

Tale of two dissimilar economies

Britain's prime minister, Mr John Major, has much to learn from President Bill Clinton's State of the Union message last week, the differences between the two economies notwithstanding.

PRESIDENT Bill Clinton has earned general acclaim for his State of the Union address last week and for its aftermath — a modest recreation of the roadshow that helped to win him the election last November. It is too soon to know for sure (and next year's budget request has not yet been published), but he may yet pull off a trick that has eluded most other elected politicians, that of persuading his electors that higher taxes are good for them. In the shadow of the US federal deficit, now more than \$300 billion a year, the proposition cannot be gainsaid even if the US Congress, although now dominated by Democrats of Clinton's party, chips away at his proposals, as much from habit as logic.

Ironically, Clinton's first diplomatic visitor of importance, the British prime minister Mr John Major, will have been able to offer only envious congratulations on last week's success. Major's embarrassment is that Britain's present problems closely resemble those that Clinton seeks to cure. There is now a comparable government deficit, rising towards 10 per cent of the gross domestic product (GDP), while Britain's deficit on overseas trade is rising steadily. Then, last week, the proportion of the British workforce on the dole rose to more than 3 million, or 10 per cent of the total. So why not, Major will have been reflecting last week, borrow a copy of Clinton's speech and arrange that something like it will be delivered by the British Chancellor of the Exchequer, who is due next month to produce a budget for the succeeding nine months? He will have decided that the remedies that promise to be effective in the United States are unlikely to be so in Britain.

Differences

Why should that be? The resemblance between the economic condition of the United States and Britain is only superficial. In both places, the recession that began more than two years ago has been sustained by the accumulation of debt on the shoulders (or around the necks) of individuals and corporations, but the burden has been lighter in the United States. Britain is also more dependent on overseas trade than the United States, and more dependent on other European economies (now also struggling to keep economic activity alive). Moreover, the British reputation for financial management has still not recovered from the disastrous events of last September, when Britain (at huge cost) was forced out of the European Exchange Rate Mechanism. It is true that interest rates have fallen, but even the reduction of the inflation rate to less than 2 per cent a year does little to

encourage people with money to spend; the depreciation of sterling by more than 20 per cent threatens that the figure will begin rising again as higher import prices begin to affect the prices British manufacturers charge for what little they still make (industrial production has fallen for two consecutive years).

Constraints

That, no doubt, is why the British government's new-style group of economic advisers told it last week (by a majority of six to one) that it should not raise taxes in next month's budget, for fear of further delaying economic recovery. The government seems ready to follow the advice, the size of its deficit notwithstanding. The snag is that there will be some point in the future, perhaps as soon as the calendar-year budget due in December, when Britain will have to grasp the nettle that Clinton seized last week. The intervening period may be hazardous. Will the government be able to borrow the funds it needs without increasing interest rates again? Or will it have to print the money and let inflation rip?

This dilemma is all too real for Major's comfort. His ministers are seeking to escape from it by looking for cuts in public spending, which by themselves would have many of the depressing effects of higher taxes, but which could usefully be used to make room for more public investment in badly needed and potentially productive infrastructure. But even if such a shift could be quickly brought about, there would remain an underlying anxiety about the future of Britain's economy. Has the capacity of British manufacturing industry (outside pharmaceuticals) to compete with the rest of the modern world been so reduced by two severe recessions in a decade that there will still be 3 million people unemployed in Britain when the present recession ends elsewhere?

Governments cannot easily admit that they have such anxieties on their minds, but others are not so restrained. And there are several gloomy straws in the wind. The old smoke-stack industries have steadily declined in Britain since the Second World War, but there has been only an inadequately compensating increase of the skill of the workers (or of their successors) thus displaced to fit them for more advanced and remunerative work. At the same time, the slew of corporate and personal bankruptcies in the past few years will have dulled the enthusiasm of this decade's entrepreneurs for following last decade's along the path of enterprise, reducing the numbers of those willing to risk their personal assets



in the pursuit of wealth-creation. Only when the recession has ended elsewhere will the British be able to estimate, from the strength of their own recovery, whether British industry can look the rest of the world in the eye.

So is there nothing that Major can learn from Clinton's example? Plenty, as it happens. Last week's speech was a clever blend of carrot (extra spending) and stick (tax increases ahead) whose persuasiveness stemmed from its detail (see page 669). Major cannot use the same recipe because his government cannot risk much extra spending or higher taxes either. But there are other ways in which a convincing message could usefully be put together. Britain, for example, now has more institutions of further and higher education than it has ever had before, and a greater need for skill than ever among the army of the unemployed. Why not subordinate for a time the cry of 'wealth-creation' to what is properly its antecedent, 'skill-creation'? And why not use the forthcoming White Paper on public research as a way of showing, not just saying, that the British government again shares with some of its predecessors the conviction that science and its intelligent application can restore the fortunes of even the most indigent economies?

But would that not just be talk, at best a promise for the future? The British, more pragmatic now than ever, have grown distrustful of promises. Nobody can blame them. But Clinton's message last week was convincing precisely because he acknowledged the seriousness of the problem of the federal deficit, the creation of his two immediate predecessors. If Major and his ministers followed suit, acknowledging that the British economy may still be on its knees when the recession elsewhere is ended, they could well discover that people would listen to them again. The belief that leaders must always express satisfaction with the state of affairs they administer is, as Clinton showed last week, an illusion.

None of that, of course, implies that Clinton is certain to succeed. The rats may get at his policy when it gets to the Congress, while nobody can be sure that he and his advisers have struck the right balance between inflation and deflation. But he deserves high marks for having identified the nature of the problem besetting the United States, and for having courageously set out to solve it. The British problem, although smaller as the numbers read, may be even less tractable. Merely facing up to it will require great daring. The best hope is that a few days in Washington will have persuaded Major that he and his ministers should try to face the downside risks of the fragile British economy — no manufacturing industry except pharmaceuticals and oil production. Then, surely, something would happen.

Meanwhile, the rest of us should not forget that while Britain, the United States and the whole of Western Europe are preoccupied with the prospect of their emergence from recession, the economic goals of just a few years ago that something substantial should be done to rescue Central Europe and parts further east from social as well as economic chaos have been relegated to a distant second place, as have the needs of the poor developing countries. Even Clinton's speech, for all its quality, had little to say on these crucial questions. □

Last-chance zoo?

The London Zoo, in the nick of time, has found a recipe that may help it to survive. Let us hope it works.

It seems at last as if there may be some light at the end of the tunnel for London Zoo after a hair-raising couple of years on the endangered species list. The zoo's £21 million plan for a "zoo of the future", unveiled last week, promises a complete financial and structural overhaul of the institution, including the installation of a new "World of Invertebrates" exhibit. Whether or not this includes an enclosure for the zoo's past directors remains to be seen. The zoo's recent troubles, mainly self-inflicted, can be explained only by a unique blend of brashness, timorousness and indifference.

This new-found perspicacity comes not a moment too soon for the ultimate owners, the Zoological Society of London, which the public was beginning to suspect of crying wolf over its financial troubles. Projected closing dates have come and gone with alarming frequency in the past two years, averted by a series of relaunches and rescue bids — including a seven-figure donation by the Emir of Kuwait in recognition of Britain's contribution to the Gulf War.

Realizing, perhaps, that military prowess can be a capricious foundation for a zoological organization, the society is sensibly using the Emir's gift as seed money for a sustained but cautious programme of fund-raising and development over the next eight years. If the zoo had shown as much prudence with the £10 million government grant it received in 1988, its survival might never have been in doubt.

The urban zoo, far from being an imperialist anachronism in these days of animal rights and the wildlife documentary, has a lot to offer the modern world, and should really have little trouble funding itself. The theory and practice of wildlife conservation, for instance, has never before attracted the level of interest and support from the general public that it does today. As people come to recognize the crucial part that zoos have to play in many species' struggle against extinction, the question of whether it can ever be ethical to keep 'wild' animals in captivity has become refreshingly less emotive.

It would also be foolish to overlook the usefulness of zoos to research. As the Human Genome Project progresses and public hysteria over gene therapy starts to abate, an accessible and comprehensive 'animal-library' looks more like a social necessity than an inhumane luxury. Once the importance of keeping animals in captivity has been accepted, the question of whether it is somehow distasteful for people to pay to see them should trouble zoos no longer.

And it is these strengths that London Zoo seems now prepared to cultivate. The new plan is still essentially a fund-raising timetable, the likes of which have been seen before, and it will take level-headedness on the part of the directors and better-conceived publicity for its goals to be met on time. But with the focus firmly on conservation and education, and with corporate sponsorship in the offing, the 1990s seem finally to have arrived in Regent's Park. □

Clinton requests more this year for NSF, advanced technologies

Washington. The economic plan that US President Bill Clinton last week presented to Congress contains a \$207-million windfall for the National Science Foundation (NSF) and \$103 million more for the National Institutes of Standards and Technology (NIST).

The president's decision to increase significantly the budgets of these relatively small agencies in the current fiscal year (which ends on 30 September) is the first tangible evidence of his support for research programmes believed likely to pay off commercially. Energy-related research and atmospheric and oceanographic studies involving supercomputers would also benefit from an attempt to provide nearly \$30 billion by the end of 1994 for a range of projects through additional federal spending and tax incentives to industry. In contrast, the stimulus package contains nothing for the \$10-billion National Institutes of Health (NIH) or for basic research in the biomedical sciences except for \$9 million to continue NIH's role in developing a high-speed computer network serving several federal agencies.

The stimulus package is part of an economic plan described in a 145-page booklet entitled *A Vision of Change for America*. Although most of the attention has gone to the new president's long-range attempt to reduce the \$350-billion annual federal deficit through a combination of tax increases and spending cuts, the short-term effect would be to restore more than half of the \$340-million increase that NSF requested for 1993 but which was rejected by Congress.

The booklet also provides details of parts of the budget for fiscal year 1994 that the president will submit to Congress on 23 March. This information, normally a closely guarded secret, is being disclosed prematurely this year because of the administration's desire to show what it is doing to stimulate the economy (see next page).

NSF intends to spend about half of its supplementary appropriation, if it is approved by Congress, on five presidential initiatives that it was not able to fund fully after its research budget for 1993 was reduced by \$13 million, to \$1.86 billion. (NSF's total budget is \$2.75 billion, which includes science education and the Antarctic programme.) Those initiatives — advanced materials and processing, biotechnology, high performance computing, global change and advanced manufacturing — were scheduled to grow by as much as 50 per cent, and the additional funds will allow NSF to come close to its original target.

Most of the rest will be spread among the

foundation's core research programmes and is expected to go towards increasing the size of existing grants and for funding requests for instrumentation. Some \$19 million has been allocated to developing software to solve computational problems relating to research on health care costs. None of the money is expected to go to science education or to the repair of academic research facilities because money for those programmes is typically stretched out over several years.

Although NSF officials consider that distribution to be fair and consistent with the president's desire to stimulate the economy quickly, many basic researchers are still smarting from reductions made in the core programme that were taken to make room for the five high-profile, interdisciplinary initiatives. They feel that NSF should stick to funding the best investigator-initiated science, a mission endorsed by the recent Commission on the Future of NSF (see *Nature* **360**, 285; 1992).

The other big scientific winner this year is the Department of Commerce, for which an additional \$262 million has been requested. The additional funds for NIST will more than double its existing Advanced Technology Program (ATP), which this year has \$68 million to spend on industrial research, by single companies or consortia, that is not as yet ready for commercialization. The programme, begun in 1990, has funded 60 projects involving an estimated \$400 million in research over the next five years. It requires companies to at least match the federal contribution.

The approach is one favoured by the Clinton administration because the ideas come from industry and undergo careful review by government scientists. By

coincidence, the fourth round of proposals were due on 24 February, and agency officials say that they could spend some of the supplementary funds, once approved by Congress, simply by going further down the list of proposals that have passed muster. A fifth round of proposals, beginning in the spring, is a distinct possibility as well.

Also within Commerce, the National Oceanic and Atmospheric Administration would get an additional \$81 million this year to buy high-performance computers for facilities such as the Forecast Systems Laboratory in Boulder, Colorado, and the National Meteorological Center in Suitland, Maryland. Another agency winner is the tiny National Telecommunications and Information Administration, which would receive \$64 million more this year to begin developing an "information highway" linking businesses, schools and government agencies. The Department of Energy would get \$47 million more to match industrial interest in the efforts of its national laboratories through so-called CRADAs (Cooperative Research and Development Agreements) with laboratory scientists.

But the rapid growth of research programmes such as ATP presents a dilemma for Clinton: although he has promised to pare the federal work force by 100,000, greater government spending requires additional staff. The NSF director, Walter Massey, has complained bitterly in recent years that the agency has insufficient resources to handle its rising work load, but the additional spending comes without any new staff positions. And although ATP officials say that the staff need not grow as rapidly as the budget, it is undeniable that peer review takes time and costs money.

Jeffrey Mervis & Tony Reichhardt

Restructuring British science

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

David Darling

Sir Graham Hills of the University of Strathclyde drew a distinction between research inputs and outputs at a meeting last week in Edinburgh cosponsored by the city's international science festival and *Nature*. Details on page 681.

Preview of 1994 US budget emphasizes big projects

Washington. A healthy increase for the National Science Foundation (NSF), delays in building the Superconducting Super Collider (SSC), more for technology, advanced computing and networks, a smaller space station and not enough for the National Institutes of Health (NIH) to keep up with inflation: that is a preview of what science can expect in the federal budget that President Bill Clinton will present next month to Congress.

The economic plan that Clinton revealed last week (see page 669) also provides an unusual advance look at the 1994 budget. Of course, there are limitations: the partial budget emphasizes projects that will stimulate the economy quickly and promote the policies of the new administration, and it is silent on many aspects of the federal research enterprise.

However, the biggest difference between the 1993 stimulus package and the forthcoming 1994 budget proposal is that Congress can enact the former without hardship while the latter must be coupled with savings of equal size. This is because Clinton has declared a temporary fiscal emergency for the current year, lifting the limits on federal spending set under a five-year budget agreement reached in 1990. But those limits will be reinstated next year.

Specifically, next month's budget request is expected to contain the following:

■ **NSF:** An overall request of \$3.1 billion, an amount not much higher than NSF sought in 1993 but almost \$400 million more than it received last autumn. As with the supplementary request, next year's budget is expected to emphasize both cooperative, cross-disciplinary initiatives and grants to individual investigators.

■ **NIH:** The \$19-billion budget for the Public Health Service, of which NIH makes up slightly more than half, is expected to increase by \$1.27 billion. But Donna Shalala, the new secretary for health and human services, has said repeatedly that she will give priority to child immunization, tuberculosis, disease prevention, AIDS treatment and women's health issues. That increase also covers more money for AIDS research and a request to inherit a \$200-million allocation for breast cancer research that this year was given to the military (see *Nature* 359, 471; 1992), leaving NIH with barely enough to match the rate of inflation, predicted to be slightly less than 3 per cent.

■ **Energy:** The administration will ask for \$640 million in 1994 for the SSC, \$123 million more than at present but \$217 million less than the Bush administration said last year was needed to keep it on target for completion in 1999. The president's science adviser, John Gibbons, said last week that there was "no reason why we have to find the Higgs boson by the turn of the century"

and that slowing down construction will give the government time to deal with the rising cost of making the accelerator's 8,600 superconducting dipole magnets. At the same time, the budget would allow Oak Ridge (Tennessee) National Laboratory to start building a \$1.7-billion Advanced Neutron Source for materials research; scientists began designing the reactor in 1984 and have tried for the past three years to get money for construction.

■ **NASA:** After daily rumours about the fate of space station Freedom, the Clinton administration announced that the station would receive its full funding of \$2.3 billion in 1994 but that much of it would be spent on designing a smaller station capable of doing less. Although congressional supporters were

Savings misnomer

Washington. A table in the appendix to President Bill Clinton's economic plan lists four ways to reduce the federal deficit. One, entitled "Managing government for cost-effectiveness and results", includes reductions in dozens of programmes that by 1997 will amount to \$35 billion. But two research projects on that list — the Superconducting Super Collider (SSC) and the space station Freedom — will actually increase government spending by billions of dollars. So why are they on the list?

The answer reflects the speed with which things can change in Washington. In the run-up to last week's speech, when the chart was compiled, those two programmes were scheduled for cancellation. Rescued at the last minute by a president worried more about jobs than about science, the programmes appear as anomalies on a list of savings. In fact, they are the only two items lacking a minus sign to denote their contribution to reducing the deficit. **J.M.**

assured that no jobs would be lost, NASA officials were last week unable to say how the station will be "restructured" and what the new version will cost. Some of the money saved is to be spent on "other NASA space missions".

■ **Commerce:** Clinton wants to double by 1998 the amount of research conducted at the National Institute of Standards and Technology's own laboratories, in addition to creating more than a hundred centres across the country to help industry modernize its production lines. The plan would nearly triple NIST's civilian budget by 1997, to \$1.2 billion. The department would continue to lead the way along an "information highway" linking businesses, schools and government agencies receiving a total of \$204 million between 1995 and 1998 to create such a network.

■ **EPA:** A new initiative in environmental engineering and technology development would receive a total of \$1.85 billion over nine years, according to the Office of Management and Budget (OMB), with a first instalment of \$36 million in 1994. Other government agencies and the private sector would be involved in this search for environmentally sound technologies and treatment techniques.

■ **Tax credits:** The president's economic plan would make permanent a tax credit for industries that increase their research spending and would give small businesses, including high-technology startup companies, a permanent tax credit for new equipment. It would also allow investors in such companies to exclude half of their gains from stock held at least five years.

Jeffrey Mervis & Tony Reichhardt

Cutting overhead costs

Washington. Tucked into the president's economic plan is a controversial proposal to lower by four percentage points the rate at which government agencies reimburse universities for the cost of carrying out research. The agencies would be able to apply the savings, estimated at \$1.2 billion over four years, to support more research, while universities would be forced to find other ways to pay for research-related expenses.

The proposal is the latest attempt by the White House Office of Management and Budget (OMB) to limit the growth of so-called indirect costs, what universities spend to administer federal research grants and to provide the necessary infrastructure. Recent investigations have found that many prominent research universities have improperly billed the government for millions of dollars. OMB would enforce the lower rate by cutting the budgets of government research agencies by an amount equal to the difference between the old and new reimbursement rates. An agency could either reduce its indirect cost payments to universities by that amount or take the savings from elsewhere in its budget.

Last month, the administration suspended a proposal to limit the administrative portion of indirect cost charges as part of a review of all pending regulations. **J.M.**

French move past Généthon to gene-therapy research

Paris. The French muscular dystrophy association (AFM) and the Centre d'Étude du Polymorphisme Humain (CEPH), which together created the Généthon laboratory in 1990, will soon part company. Généthon will stop its mapping programme this summer after a run of scientific successes, including a 'second-generation' genetic map of the human genome and the physical map of chromosome 21 (see *Nature* **359**, 380 & 794; 1992), and AFM's substantial resources will subsequently go towards eliminating bottlenecks in gene-therapy research. Daniel Cohen will return to his roots at CEPH, where he will build the sequel to Généthon, an industrial-scale gene discovery centre (see *Nature* **359**, 526; 1992).

Cohen says that Généthon will have finished crude maps of the genome by June and that the work will be continued by genome centres worldwide, including two in the United States that he helped to create. "We have accomplished what we set out to do, this is the end of the game", he says.

But Piotr Slonimski, head of the new French genome agency, is reluctant to abandon an area in which France has excelled and he says that much remains to be done. He wants to maintain the mapping centre, but says that the agency has no plans as yet to pay for the cost of operating Généthon. Another possibility is that research groups elsewhere might pay for use of the centre.

In shifting its support to gene therapy, AFM has no illusions of repeating its spectacular foray into genome research. It recognizes that it cannot support the enormous cost of research, even for a single disease, and that success is elusive. Instead, AFM will combine its \$16.9 million annual contribution to Généthon with money from other charities to help researchers address the major structural drawbacks to gene therapy research in France.

"What we need most are more sorts of vectors and the large amounts required for tests in animals and humans", says Axel Kahn, director of the INSERM Laboratory of Research on Genetics and Molecular Pathology at the Cochin Institute of Molecular Genetics in Paris.

AFM's solution is to create a host of biotechnology companies at the Généthon site in Ivry outside Paris and to guarantee their profitability by generous subsidies. Kahn says such companies would relieve researchers of much of the laborious work in producing new constructs; he believes that US scientists have made so much progress because of the existence of such companies as Somatic Therapy, Genetic Therapy and Viagene. There are no such companies in Europe, let alone in France. AFM will also

create a new laboratory at the Ivry site dedicated to testing the effect of adding unknown genes to mice and to creating animal models for particular diseases.

Cohen agrees that it is more appropriate for medical charities to ease scientific bottlenecks than to fund long-term research programmes, which he sees as the government's responsibility. He says he has always encouraged AFM to "strengthen the scientific community and promote the scaling-up of biology research". But although he approves of the new direction, he says that "it's not for me" because it would not suit his preference for industrial-scale research.

Instead, Cohen will return to CEPH in central Paris, which he created in 1982 with Jean Dausset, a Nobel prizewinner in medicine. Cohen intends to expand CEPH and to double its budget of FF40 million (US\$8 million) by the end of the year through the recently created Jean Dausset Foundation, which has given him FF100 million to develop a 3,000-square-metre laboratory at CEPH.

Cohen says he will build an industrial-scale centre, bigger than Généthon and using prototype technology, to discover genes involved in multigenic diseases such as cancer and cardiovascular disorders. He also hopes to attract other companies to the site.

The move towards gene-therapy research in France coincides with the expected first approval within the next few weeks of a gene-therapy trial involving patients with adenosine deaminase (ADA) deficiency. This will be carried out by Alain Fischer, who

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Daniel Cohen outside Généthon

heads France's largest gene therapy group at the Hôpital Necker in Paris, in collaboration with Dinko Valerio in the Netherlands (see *Nature* **355**, 190; 1992).

Although gene-therapy trials were approved in 1990 when the national bioethics advisory committee declared that somatic gene therapy posed no special problems, Fischer found the approval procedure more difficult than he anticipated. The trial was approved as safe by the state genetic engineering commission but had to go before an additional internal INSERM committee to meet a requirement that groups be insured against any adverse consequences. Moreover, Fischer says that the local ethics committee has dragged its feet because of its reluctance to take risks in the wake of the French blood contamination scandal (see *Nature* **359**, 764; 1992).

Declan Butler

London meetings on British science

On Friday 19 March, there will be a meeting at the Royal Society, 6 Carlton House Terrace, London SW1, to discuss proposals for the forthcoming White Paper on the organization of British science. The speakers will be Sir Eric Ash (Rector, Imperial College London), Professor Michael Brady (University of Oxford), Dr Dai Rees (Secretary, Medical Research Council) and Sir Mark Richmond (Chairman, Science and Engineering Research Council); the meeting will start at 9.30 a.m. and end at 4.00 p.m.

Admission to the meeting will be by ticket, free of charge, obtainable from Mary Sheehan, *Nature*, 4 Little Essex Street, London WC2R 3LF. Coffee and tea will be provided.

There will also be a sandwich lunch for those who want it, **at a cost of £5**: please send a cheque made out to *Nature* with your ticket application. Tickets for both the meeting and lunch will be sent out during the second week of March.

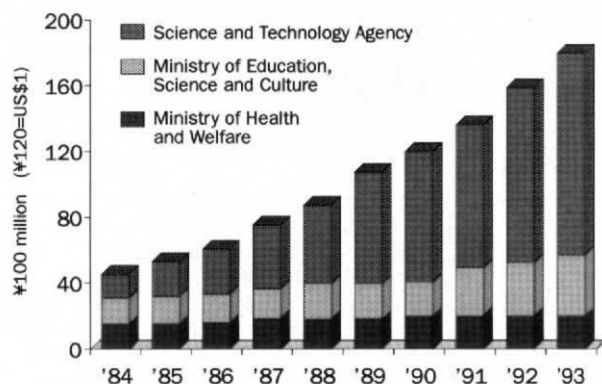
* **On Thursday 18 March**, Save British Science is holding a meeting called "Waldegrave meets the scientists" at 12.30 p.m. in the Great Hall, Sheffield Building, Imperial College, London SW 7. The Rt Hon. William Waldegrave, MP, Chancellor of the Duchy of Lancaster and Minister of Public Service and Science, will speak and will answer questions. Entry free. Seats may be reserved by application to Save British Science, PO Box 241, Oxford OX1 3QQ (tel. 0865 273407, fax 511370).

Japanese panel proposes extension of cancer research

Tokyo. One of the most important sources of funds for cancer research and biological research in Japan seems likely to keep flowing for another decade. Earlier this month, a government working group of scientists with the backing of the Ministry of Health and

Nakasone. The joint funding by MHW, MESC and STA is unusual for Japan, where government ministries and agencies tend jealously to guard their turf. By March 1994 the three agencies will have spent more than ¥100 billion (US\$900 million).

What Japan spends on cancer research



Welfare (MHW), the Ministry of Education, Science and Culture (MESC) and the Science and Technology Agency (STA) recommended the extension for another ten years of the Comprehensive Ten-Year Strategy for Cancer Control, with a slight shift towards more practical applications.

The cancer programme was launched in 1984 by then Prime Minister Yasuhiro

topics (see table). New among them are research into cancer prevention and ways to improve the quality of life of cancer patients.

In line with the wishes of biologists working under the education ministry, the topics also include cell biology, immunology, virology and molecular research. Tadatsugu Taniguchi of Osaka University's

About half of the budget has been spent by STA on a heavy-particle accelerator for cancer treatment. The programme has also provided a significant boost in research funds not only for cancer researchers working for MHW but also for molecular biologists, cell biologists, virologists and immunologists in Japan's universities, who are supported by MESC.

The working group, headed by Masaaki Terada, director of the National Cancer Centre Research Institute in Tokyo, recommends that the next phase should focus on seven

An agenda for cancer research

- Molecular mechanisms, including:
 - oncogenes, tumour suppressor genes
 - radiation-induced cancer
- Metastasis/invasion and characterization
 - mechanism and control
 - abnormalities in expression, cell death
- Predisposition and immunology
 - familial incidence
- Prevention, including:
 - risk factors, epidemiology
 - education
- New diagnostics
 - imaging, early diagnosis
- New treatments
 - gene therapy, drug tolerance
 - heavy-particle beam
- Quality of life
 - after-care, reducing pain
 - economic factors

Institute for Molecular and Cellular Biology, an important figure in the present programme, says that he expects the programme to "continue to feed Japanese biology in general" because of its contribution to better understanding of the disease.

On the other hand, Yusuke Nakamura, chief of the division of biochemistry in the Cancer Institute, Tokyo, whose group of 40 researchers is expected to play an important role in the development of new diagnostics, says that his chief motivation is to "save patients' lives". He hopes to develop DNA diagnostics both for early diagnosis of cancer and to assess the degree of malignancy of tumours so that unnecessary surgery can be avoided.

The working group also called for research on gene therapy. "Many clinicians are interested", says Terada, but few Japanese scientists have research experience in gene therapy. The Japanese government has only recently initiated a process to develop guidelines and approval processes for such research, which for the foreseeable future will be confined to animals. The present ten-year programme also led to the establishment of the first postdoctoral positions supported by the Ministry of Health and Welfare, and it also funds a small exchange programme between Japanese and foreign scientists.

Although the programme's budget has yet to be determined, the working group hopes that research funds will be increased by about 50 per cent. Its report will undergo minor rewriting by its parent committee before approval is sought from the cabinet in time for the three agencies to submit their budget requests at the end of August.

David Swinbanks

Stretching the truth on cancer

Tokyo. The Japanese Science and Technology Agency and the ministries of Health and Welfare and Education, Science and Culture have gone too far in extolling the achievements of their ten-year anti-cancer project.

A 12-page colour pamphlet issued by the three government organizations describes for the public an impressive list of accomplishments. Among them are the discovery that the HTLV-1 virus causes adult T-cell leukaemia (ATL) and that the hepatitis-C virus is the cause of many cases of liver cancer. The pamphlet also includes an illustration of the drug Camptothecin under the heading "new anti-cancer drug treatments developed".

The problem is that the claims are not true. The link between HTLV-1 and ATL was established by Japanese and US researchers before the start of the project. The hepatitis-C virus was first isolated and characterized by Chiron, a US company, and Camptothecin was independently developed by a Japanese company, Yakult.

Masaki Mugitani, deputy director of the disease control division of the Ministry of Health and Welfare, says that the exaggerations are a result of writing for a general audience. A more technical report by and for scientists, he says, explains that the project helped to elucidate the mechanism by which HTLV-1 causes cancer and the route of infection and that it led to the discovery and characterization of the Japanese strain of hepatitis-C virus.

The pamphlet is an embarrassment, says Masaaki Terada, director of the National Cancer Center Research Institute and leader of the working group that reviewed the project. It will be withdrawn and reissued after revision.

D.S.

Leprosy researchers lament suppression of Indian paper

New Delhi & Washington. Professional jealousy has for more than a decade blocked publication of a paper on leprosy that could have hastened progress on finding an animal model indigenous to India, where the disease is most prevalent.

condemned by the scientific community and because they are saddened by the loss of an important animal model for leprosy research. There is also the matter of professional pride. "It was in our grasp to be the first to receive credit [for documenting experimental leprosy in a colony of some 100 lorises inoculated with *Mycobacterium leprae*]", says one scientist.

Gerald Cusick/Bruce Coleman Ltd.

The work did not advance to the stage where vaccines against the disease could be tested. "The loris was a new experimental animal and a lot of basic work had to be done before inoculation could begin", says Bedi, the principal investigator of the US-Indian project for several years and now a skin specialist in a Delhi hospital. The published work was to be based on observations over a few years of a spreading form of leprosy.

But the observations lasted for only six months. Agarwal, who was about to leave the institute, assumed control of the project from Bedi's successor, M.N. Ghosh, and killed the 66 surviving animals. "Agarwal destroyed valuable evidence of the

progressive development of the disease in loris in order to grab credit for himself before he left", says one researcher who requested anonymity.

Agarwal denies that he acted improperly and says that he deserves to be listed as first author because he was the principal investigator when the research ended and because he was the one to detect leprosy transmis-

sion. "If these lorises had not been detected by me, there would have been no findings, no publication and no controversy". He also says that he killed only five of the research monkeys and that the rest died of natural causes.

Kirchheimer, now retired, disputes Agarwal's account, saying that Agarwal "knew nothing about leprosy" and "had no interest in the disease". He says that Agarwal acted out of "jealousy and disappointment" after failing to succeed M. Balasubramanian as director of JIPMER.

Curiously, Kirchheimer was himself involved in a celebrated dispute in the late 1970s over the discovery of several natural cases of leprosy in armadillos in the southwestern United States. Kirchheimer accused scientists at a research institute in Louisiana of allowing infected animals to escape and of causing the artificial spread of the disease. Although that charge has since been refuted, Kirchheimer remains critical of what he sees as widespread unethical conduct among some of his former colleagues.

One veteran research worker who was surprised to hear about the Indian research was Robert Gormas, who has worked at the Delta Primate Center in Louisiana for 14 years, successfully transmitting the disease to dozens of sooty mangabees and rhesus monkeys. More recently, he has discovered two natural cases of leprosy in sooty mangabees.

He says that Bedi's work "could have been used widely" if it showed that the slender loris was a useful model. Wayne Meyers, of the US Air Force Institute of Pathology in Washington, DC, says that researchers "are still looking for an animal model that is easily maintained, readily accessible and relatively inexpensive" and that it would be a boon to Indian researchers if the loris filled the bill.

K.S. Jayaraman & Jeffrey Mervis

The slender loris is common in southern India.

A group of Indian and US scientists working at JIPMER (the Jawaharlal Nehru Institute for Postgraduate Medical Education and Research) in Pondicherry found that leprosy could be experimentally transmitted to the slender loris, a species of monkey easily obtainable in southern India. At the time researchers were looking for an alternative to the armadillo, which is expensive and not native to Asia.

The 1981 article was accepted by the *International Journal of Leprosy* pending resolution of a dispute over authorship. An influential microbiologist at the institute, S.C. Agarwal, insisted on being listed first; his co-authors, including Waldemar Kirchheimer, then of the US Public Health Service hospital in Carville, Louisiana, and B.M.S. Bedi and E. Narayanan of JIPMER, objected on the grounds that Agarwal had not been involved in the experiment.

The journal's editor, Robert Hastings, says that he asked Agarwal to "clear up the flap" but that Agarwal never replied. Three years later, Agarwal refused to allow another scientist to present the results at a World Leprosy Congress in Agra, India, and to this day they have remained unpublished.

Agarwal's former colleagues say that they have finally decided to speak up because such conduct should be exposed and

McBride found guilty of fraud

Sydney. William McBride, the Australian researcher credited with issuing the first warning of the dangers of the drug Thalidomide, has been declared guilty of scientific fraud.

The medical tribunal of the state of New South Wales last week found McBride guilty of falsifying the results of an experiment to establish whether the morning sickness drug Denebox, produced by Merrell Dow, causes birth deformities. The tribunal dismissed all but one of a number of charges relating to the separate issue of McBride's conduct as a gynaecologist, and described that one proven charge as a "minor lapse". The tribunal will determine a penalty after McBride's defence lawyers and the state health department prosecutors have studied the lengthy judgement.

An inquiry set up in 1988 by the research centre founded by McBride, Foundation 41, found that he had altered the results of an otherwise inconclusive experiment on 16 rabbits. The inquiry found that McBride had changed the drug doses, added rabbits and a control group and identified non-existent abnormalities in the rabbit fetuses in an attempt to show that the drug is dangerous.

The case has forced researchers and government officials to pay closer attention to the issue of scientific misconduct. The Australian National Health and Medical Research Council now requires institutions applying for grants to have procedures in place for dealing with allegations of scientific misconduct.

Mark Lawson

Britain wins bid to host ocean drilling office

UK oceanographers ride the waves of uncertain funding

London. British marine geophysicists are rejoicing over the successful bid by the University of Wales, Cardiff, and the Natural Environment Research Council to be the

Ocean Drilling Program

first non-US host of the science planning office of the international ocean-drilling programme (ODP).

The National Science Foundation, which runs the programme, provides just over half of its £28 million (US\$40 million) annual budget. The rest comes from its six

other members, including four individual countries — Britain, France, Germany and Japan — a consortium of smaller European countries organised through the European Science Foundation and a joint membership of Canada and Australia (Russia has withdrawn because of financial difficulties).

The scientific content of the ODP programme, begun in 1985, is coordinated through a body known as the Joint Oceanographic Institutions Deep Earth Sampling (JOIDES). Subscriptions from members are used to pay for a converted oil explorations

ship, the JOIDES *Resolution*, based in Austin, Texas, under contract to JOI, Inc., a Washington-based organization.

The JOIDES planning office, which coordinates the programme's scientific agenda, coordinates the work of a number of committees. It has been based for successive two-year periods at a series of US universities; its home until the end of 1994 is the University of Washington in Seattle.

It was decided to locate the office outside the United States to encourage scientists from other countries to feel more involved in planning the scientific agenda. Cardiff won the contract, worth £400,000 (US\$600,000) for two years, against two other international bids.

A major factor behind JOIDES's decision to locate its planning office in Cardiff is the return of the JOIDES *Resolution* later this year from the North Pacific to the North Atlantic, where it carried out its first drilling in 1985. "People beat a path to your door to tell you what is going on", says Joe Cann, professor of earth sciences at the University of Leeds. "Having a pipeline of hot topics coming into Cardiff gives us a big scientific advantage in planning our own research." Projects planned for 1994–95 include investigations of the hydrothermal mounds on the mid-Atlantic ridge and a drilling programme on the Barbados ridge to look at fluid flow in ocean-bed sediments.

Two other ODP contracts are also open for bids. One is for a new computing system on the ship, and the other for the data storage — or "logging" — facilities now provided by the Lamont Doherty Observatory in New York.

David Dickson

London. Recent funding for British oceanographic science would mirror the topography of the subject — a series of peaks and troughs.

The good news is that the Natural Environment Research Council (NERC), which finances most of Britain's research on the oceans, has found the money for some badly needed equipment for its fleet of research vessels and that the University of Wales, Cardiff, has been picked to host the JOIDES planning office (see left). The bad news is that the NERC has been forced to curtail the number of research cruises planned for next year. In particular, the *RRS Challenger* will be laid up for one month in three so that only one crew is needed, and its sister ship, the *RRS Darwin*, is scheduled to be used two months for charter even though there are research projects that could have filled the slot.

Two factors have affected NERC's planning. The first has been the general squeeze on government funding for research. The second has been an unexpected increase in the cost of moving the research vessel service from its current base in Barry in South Wales to the Centre for Deep-Sea Oceanography in Southampton.

The resulting budgetary restrictions have meant a reduced investment in scientific equipment on the various research vessels, including the *RRS Darwin*, which was built in the mid-1980s, and the *RRS Discovery*, which has recently undergone a major conversion.

Some of these concerns have now been assuaged. Last year's decision to purchase a swath bathymetry system was followed at the beginning of this month by an announcement that an extra £400,000 being made available through a reduced subscription to CERN (see *Nature* **361**, 485; 1993) is to be used to buy a second TOBI (Towed Ocean Bottom Instrument) sidescan sonar package developed by the Institute of Oceanographic Sciences. In addition, the NERC is reported to be about to purchase for the *Discovery* two air compressors, essential equipment for profiling the ocean bed.

But there is still concern that cost reductions will prevent scientists from making full use of the research vessels. Further cuts in the Darwin's scientific cruises are threatened if the NERC fails to rent out the vessel in February and March 1994. And the opportunities for research on foreign-owned vessels have been reduced by the NERC's reported decision to shelve a scheme under which such facilities have been made available in return for offering foreign researchers access to British cruises.

David Dickson

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

The JOIDES Resolution

French AIDS team wins Faisal Prize

London. Three French AIDS researchers and two physicists, one German and one from the United States, have received the 1993 King Faisal International Prize in medicine and physics respectively. The prize is also awarded annually for service to Islam, Islamic studies and Arabic literature, and each recipient receives \$100,000.

This year's medicine prize honours Luc Montagnier, Jean-Claude Chermann and Françoise Barre-Sinoussi for their work on identifying HIV and the way in which it damages the human immune system. Working together in the viral oncology unit of the Pasteur Institute in Paris, the three researchers isolated in 1983 what they then described as the "lymphadenopathy-associated virus" (LAV), subsequently shown to be the virus responsible for AIDS.

Montagnier has recently set up a new body known as the World Foundation for Research and Prevention of AIDS (see *Nature* **361**, 287; 1993). Barre-Sinoussi is head of the Retrovirus Biology Laboratory of the National Institute of Health and Medical Research (INSERM), based at the Pasteur Institute. Chermann is now director of INSERM's Laboratory of Retrovirus and Associated Diseases in Marseille.

The physics prize goes to Herbert Walther, director of the Max Planck Institute for Quantum Optics, and Steven Chu, chairman of the department of physics at Stanford University, for their work in quantum optics. Walther invented the technique of trapping single photons in electromagnetic fields and Chu has developed new approaches to studying such systems.

David Dickson

Global project under way to sample genetic diversity

Washington. An international group of population geneticists, molecular biologists and anthropologists want to sample the world's population for genetic diversity — if they can find \$25 million to do the job properly.

The Human Genome Diversity Project hopes to use the diversity of living people to understand human evolution. The organizers say that the need is urgent because increasing urbanization threatens the integrity of small isolated populations on which this analysis depends. The information should also prove useful to the Human Genome Project, which does not consider population-level variation.

Ideally, the project would create thousands of cell lines from the nuclear DNA of sampled individuals. As a potentially permanent source of DNA, these lines could be used to answer questions about ancient migration patterns and the relative importance of selection and drift in divergent evolution as well as to improve understanding of the distribution of disease susceptibility or resistance.

The project must balance its desire for information against its limited resources. For example, the best sources of nuclear DNA, cell lines transformed by the use of blood samples, are also the most expensive and require laboratory procedures within 72 hours of sample collection. Although building new facilities and transferring technology are part of the general plan, where and how such laboratories can be built will limit the populations that can be sampled. The

organizers are also hoping that newer technology will provide cheaper and faster ways of obtaining data, such as using cheek swabs or hair follicles instead of blood to yield transformed cell lines.

Carrying out uniform DNA analysis, another priority for the project, may also require new technology if older data are to be made compatible with those collected using simpler markers and procedures. There are also ethical issues, including whether samples should be screened for pathogens such as HIV and how to prevent the misuse of information on diversity.

The project, with a steering committee of five European and five US scientists and Luca Cavalli Sforza of Stanford University as chairman, has sponsored three workshops since last July to discuss sampling techniques, ethical issues and questions about funding. The committee hopes to present its recommendations at a fourth and final meeting next September in Alghero, Sardinia.

Population geneticists and statisticians in the first workshop set a lower limit of sampling at 25 unrelated individuals from 400 populations, recommending that 100 loci be analysed in describing the genome. At the second workshop, anthropologists compiled lists of priority populations for each region using groups called Isolates of Historical Interest (IHIs). These small populations are more likely to reflect direct descent from earlier small groups: only inbred groups differ from each other as a result of

divergent evolution. As individuals mix with those in urban centres, evolutionary data are lost. Target populations were also picked to represent all major language groups.

The organizers have prepared a tentative aggregated list of priority populations including both IHIs and their nearest neighbours, which they hope to make public for comment. The list is not complete because all regions were not equally represented at the meeting and many scattered populations have already been sampled. For example, an international project to study the HLA locus has for 25 years been creating cell lines in 400 labs around the world.

Prospects for obtaining the estimated \$25 million needed over the next five years are uncertain. Meeting last week at the US National Institutes of Health (NIH), the committee talked to representatives from NIH, the National Science Foundation and the Department of Energy, who offered encouragement but no promise of funding. The group is also hoping to gain a place in a five-year plan being written now for the Human Genome Project.

Money is being sought in the next Framework programme of the European Communities for eleven laboratories to draw samples throughout the Middle East and Eastern Europe, and funding agencies in China, Japan and India have expressed support for sampling projects in their own countries. In the meantime, investigators are using money from existing grants for small-scale sampling.

Jenna Roberts

Bart's loses fight against merger

London. St Bartholomew's Hospital in London — known as Bart's — has reluctantly accepted the conclusions of a report commissioned by Britain's Department of Health that it should merge with the nearby Royal London Hospital (see *Nature* **361**, 194; 1993).

Its decision came after the Health Minister, Virginia Bottomley, announced that she had accepted the report's conclusions that London's teaching hospitals must be merged because of declining demand for hospital services. But Bart's still hopes to retain its historic Smithfield site, which it has occupied for 900 years.

Bottomley also announced a reprieve for two smaller specialist hospitals, the Royal Marsden Cancer Hospital and the Royal Brompton National Heart and Lung Hospitals. Both had protested that their research activities could suffer if they were required to merge with the Charing Cross Hospital, as the report had suggested. Special reviews will now be carried out of facilities for treatment and research in six specialist areas in the capital. These reviews will be completed by the end of May.

David Dickson

Italy tries to define basic research

Munich. The Italian Ministry for Universities and Research has sent a scientific task force into its space agency ASI to sort out "as a matter of urgency" the continuing deadlock in the funding of basic space science. The experts have a month to decide how much of ASI's IL800 billion (US\$520 million) annual budget should go to fundamental research, resolving a prolonged and bitter feud between the agency and its scientific advisers that has blocked the flow of billions of lira to scientists.

Italy's scientific community initially welcomed the creation of ASI in 1988 to coordinate national space research. But continuing political infighting has hindered its work and led to much criticism of its conduct. Most damaging has been the disagreement about interpretation of the law requiring ASI to spend at least 15 per cent of its budget on basic research. Ambiguous wording of the phrase led to different

interpretations: ASI says that the Italian money spent by the European Space Agency (ESA) on construction of platforms should be included, while its scientific advisory committee believes that such spending is really applied. The difference amounts to billions of lira. The task force, which includes ESA president Francesco Carassa, has been asked to decide the question.

The dispute has delayed funding by several months and has opened up serious rifts within the advisory committee, which approves grant applications. The committee last September voted out its chairman, Remo Ruffini, who refused to approve any funds until the issue was resolved. Despite its sympathy to the cause, the committee felt that this unilateral action was detrimental to Italian research. But science minister Antonio Ruberti sided with Ruffini and the committee has not met since then.

Alison Abbott

Help needed for Eastern Europe

SIR — I am a Hungarian researcher visiting the National Cancer Institute in the United States. I am pleased to learn (*Nature* 360, 504; 1992) that the Howard Hughes Medical Institute "is weighing the idea of supporting researchers in the republics of the former Soviet Union and Eastern Europe". Researchers who receive such grants will truly benefit. But the contrast between the amount of individual grants in the States and those being considered for scientists in Eastern Europe and Russia is very disappointing.

We in Eastern Europe are really in desperate need of funds for research. However, my three-year research experience in Austrian, UK and US laboratories and my ten-year experience in Hungary in the field of environmental carcinogenesis make me say that the sums being considered for individual grants are so low that they will hardly provide any help to any scientist in the region. Laboratory instruments, chemicals and small equipment cost the same, usually even more for us than for institutes in the West, probably because the market is much smaller in Eastern Europe. The only components of research that are cheaper there, sadly enough, are manpower and expertise. No wonder the Howard Hughes Medical Institute is cautious when planning expansion in that part of the world, but in my view, instead of "supporting" many projects with irrationally low amounts of money, fewer but well-selected and well-funded grants would provide more benefit for the recipients and for the scientific progress of these countries.

Bernadette Schoket

Department of Biochemistry,
National Institute of Public Health,
Gyali ut 2-6,
Budapest, H-1097 Hungary

More interferons

SIR — While writing a review on animal interferons, I found that a computer search of the literature identified 16 papers that referred to interferon-tau. These papers described many interesting features of this molecule which appeared to be identical to interferon gamma and has been identified in humans, mice and swine. My initial assumption was that this was a new nomenclature reflecting the molecule's T-cell origin. But I was unable to find a citation referencing this nomenclature or any relevant statement from the International Interferon Nomenclature Committee.

Another reading of some of the pap-

ers, however, pointed to an alternative explanation — a persistent typographical error. When working in some word-processing programs, one is obliged to use ASCII codes for Greek letters. Unfortunately, there is no ASCII code for gamma, a source of enormous frustration to immunologists. The letter tau (τ) however, bears some resemblance to gamma (γ) and could be inadvertently substituted. Indeed, in the past I have used tau as a temporary substitute for gamma in a rough draft. A call to an author of one of the papers confirmed indeed that the use of tau was a typographical error. The substitution of tau for gamma clearly occurs with sufficient frequency that a respectable literature has built up on the subject of interferon-tau.

The failure of our educational system to provide a classical education and the uncritical use of the word processor thus appear to have contributed to one of the major problems in modern immunology, the explosive growth in the number of cytokines. It also raises the possibility that other cytokines or peptide chains do not exist except as persistent typographical errors.

Ian Tizzard

Department of Veterinary Pathobiology,
Texas Veterinary Medical Center,
Texas A&M University,
College Station, Texas 77843-4467, USA

Diet and cancer

SIR — In the leading article on diet and breast cancer (*Nature* 359, 760; 1992) the author's conclusion, based on a single cohort study, that the fat-breast cancer hypothesis should be consigned to oblivion, is premature, to say the least. The association between dietary fat and breast cancer was first demonstrated in animal models more than 50 years ago. Since then, numerous studies, using chemical, viral and X-ray-induced mammary tumour models, have shown that high fat intake promotes mammary cancer. In addition, cross-country comparisons and studies of migrants from low to high risk countries support the fat hypothesis. Data from case-control and cohort studies (such as the Harvard Nurses Study) have yielded much less consistent results. One possible reason is that the range of fat intake within the US population is too narrow to make any difference in cancer risk.

Evidence from China, Japan and animal models suggests that protection occurs if the diet contains 20 per cent fat calories or less. If this is the case, the results of the Harvard study could be interpreted to mean that even the lowest quintile of fat intake in the US female population (27 per cent) is above the

hypothetical threshold where dietary fat ceases to exert tumour-enhancing effects. Several dietary intervention trials in women at high risk for breast cancer are under way in Canada, Sweden and the United States to test this hypothesis. Feasibility studies already completed have shown that the stated goal of these trials, that is, reduction of fat intake to 20 per cent of total calories, can be achieved and maintained over a two-year period. Against this background, one would be hard pressed to conclude that the fat hypothesis is "little more than wishful thinking", particularly on the basis of a single, albeit large, cohort study.

Leonard A. Cohen

Division of Nutrition and Endocrinology,
American Health Foundation,
Dana Road,
Valhalla, New York 10595, USA

Fraud inquiries

SIR — As officers of the Association of Researchers in Medicine and Science, we read with interest the recommendation of the inquiry into scientific misconduct at Michigan State University (*Nature* 361, 287; 1993) that the university "should develop policies for resolving conflicts . . . for handling allegations of scientific misconduct".

When one of our members recently obtained evidence that a graduate student had been fraudulently altering data and the head of department sided with the student without considering the evidence, it became clear to us that there exists no accepted mechanism in UK universities through which a scientist can appeal for arbitration, or indeed for help or advice, in a dispute of this kind. It would clearly be slow, costly and unnecessarily cruel to follow the American pattern.

Our association, which is concerned about the quality of scientific research, proposes that British universities should set up an independent body of scientists to which involved scientists could appeal. A possible model would be the General Medical Council except that it would be concerned solely with complaints made by scientists themselves. Members of professional societies, the Royal Society and editors of scientific journals could provide names in each case of suitable people who are properly independent. It would seem wise for the body to sit in private, to protect the interests of both whistleblower and accused and to avoid the expense of lawyers.

Peter N. Campbell

Janice Taverne
University College London,
London WC1E 6BT, UK

The genesis of the UK coal crisis

Mike Parker and John Surrey

The UK government finds itself in a tangle over energy policy. That reflects the Conservative party's own agenda in dealing with the coal and nuclear industries since it came to power in 1979.

THE announcement on 13 October 1992 that 31 of Britain's 50 remaining deep coal mines would close forthwith, with 30,000 redundancies, led to a wholly unexpected political storm, followed by a stay of execution for 21 of the 31 pits, and an undertaking by the President of the Board of Trade to review the prospects of the 31 pits in the context of energy policy (see ref. 1). The immediate factors behind the announcement to close the 31 pits were the near certainty of a steep fall in coal demand when the existing power-station coal contracts end in March 1993, and the government's determination to adhere to its timetable for privatizing the nationalized coal industry, which appeared then to preclude a phased closure programme.

Whatever decisions flow from the government's policy review, it is important to understand the background to the coal crisis. Here we examine the objectives and results of government policies for coal and nuclear power since the first Conservative government under Margaret Thatcher entered office in 1979. During this period there was a sharp transition from high to low energy prices in the mid-1980s, with a decline of traditional concerns over the security of energy supply. Towards the end of this period, environmental concerns became much more important and the electricity supply industry was privatized, with adverse consequences for both coal and nuclear power. Throughout, there was no formal framework of energy policy.

Coal policy

The government's objectives for coal emerged shortly after the 1979 general election and were held consistently over the whole period until 1992. They were aimed at breaking the power of the National Union of Mineworkers (NUM), identified as the cause of the downfall of the previous Conservative government in 1974 and subsequently the humiliating climb-down over pit closures in February 1981, and demonstrating the superiority of private enterprise over public ownership. As Nigel Lawson, later Secretary of State for Energy and then Chancellor of the Exchequer, wrote in April 1981, "Our original aim was to build a successful, profitable coal industry independent of government subsidies, to demonopolize it and ultimately to open it to private enterprise. . . . Then the events of

February 1981 showed beyond reasonable doubt that we will make no progress towards our aim until we deal with the problem of monopoly union power"². These objectives required the contraction of the industry as far and as rapidly as practical considerations allowed.

The main instrument of the policy was financial discipline applied steadily (apart from February 1981) in a manner compatible with the problem of NUM opposition to pit closures until the 1984–85 miners' strike, and thereafter with constant pressure to rationalize output and reduce costs. The government also avoided output targets (except informally for profitable open cast production) or the use of long-term energy forecasts to justify particular levels of output. Almost all the large costs of miners' redundancy payments and other restructuring costs were met by the government to facilitate the rundown.

The political agenda and financial discipline over-rode traditional 'energy policy' considerations. Even the pro-coal statements in 1979–80 during the OPEC II oil crisis were largely rhetorical. Because of its importance and nature, the political agenda on coal was low-key and covert, with very little attempt at a publicly justified rationale. Electricity privatization, involving the division of the Central Electricity Generating Board (CEGB) into two generator companies (PowerGen and National Power)

assisted the policy objectives on coal. The government encouraged the development of competition in power generation, which involved introducing large numbers of new combined cycle gas turbine stations (CCGTs): the so-called "dash for gas". By late 1992, CCGTs built or under construction have preempted nearly half the market for British coal in power generation in the mid-1990s. The excessive market power of the two privatized generators, the many separate companies within the privatized electricity industry, the form of price regulation, and the build-up of generators' coal stocks, all made it impossible to arrange new contracts which would avoid large volume and price reductions for British coal following the end of the initial 3-year coal contracts in March 1993. Throughout, the government showed indifference to the adverse effects of electricity privatization on the coal industry, of which it was warned four times by the House of Commons Select Energy Committee from late 1988 to early 1992. After the 1992 election, the government was working to a timetable to privatize British Coal as soon as possible, based on 'freely negotiated' contracts with the generators. This was to be the "ultimate privatisation", pledged in October 1988 by Cecil Parkinson, then Energy Secretary, who later explained that "What was ultimate about the proposed privatisation of coal was

PA

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

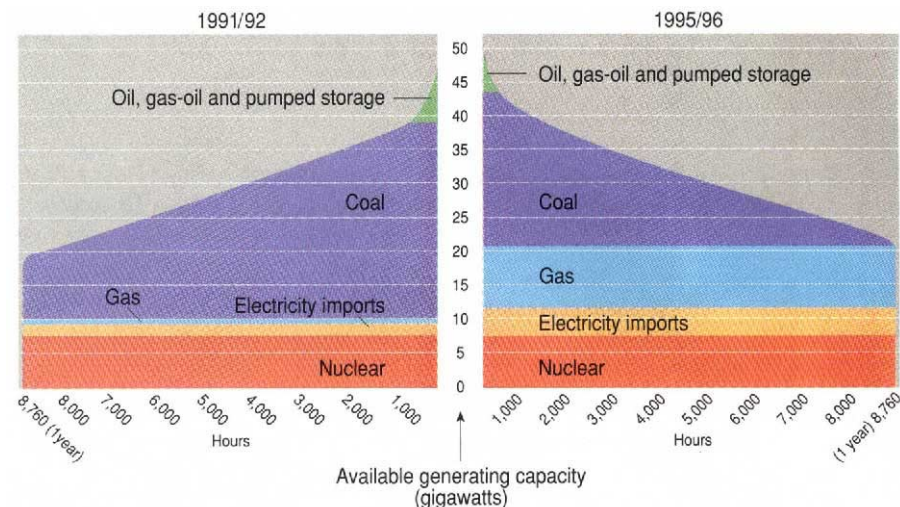
Public protests — demonstrators at a miners' rally, London, October 1992.

that it would mark the end of the political power of the National Union of Mine-workers and would make the coal industry what it should always have been, another important industry, no more and no less important than many others"³.

Until October 1992, the political objectives seemed to have been largely achieved. A greatly slimmed-down industry was shortly to be privatized, and the NUM had lost its power. This was because, first, Conservative victories in the preceding four general elections allowed the policy to be pursued consistently over a long period. Second, external market forces and environmental developments since the mid-1980s had reinforced the political agenda, so that policy went with the grain of events. From the mid-1980s onwards the policy could be justified in terms of low coal import prices and environmental issues of public concern, namely 'acid rain' and the 'greenhouse effect'. Third, there had been careful choice of appropriate policy instruments, flexibly and cautiously applied (except in February 1981), with no fundamental conflict with Treasury or other policy objectives. Fourth, the government was fortunate in having Arthur Scargill, leader of the NUM, as its opponent in the 1984-85 strike, since his actions led to the creation of the politically moderate Union of Democratic Mineworkers (UDM) and the continued working of the UDM pits in Nottinghamshire, without which the result of the strike might well have been different. Finally, unlike in Germany, the coal industry had no political constituency of any real influence. Apart from a few Conservative members of parliament representing 'UDM' constituencies, the support was concentrated in opposition-controlled areas. With its large majorities in the House of Commons, the government was immune to political pressure on behalf of the coal industry and the miners.

Nuclear power policy

The decision in November 1979 to launch a pressurized water reactor (PWR) programme, when oil and other fuel prices were at their peak, appeared to provide the direction and political will necessary for the long-term development of nuclear power in the United Kingdom. But appearances were deceptive. The House of Commons Energy Committee showed that the foundations of the enterprise were shaky — in particular, the absence of capacity need, the sensitivity of the economics to small changes in the key assumptions, and the doubts about the future performance of the nuclear plant industry and the public acceptability of a PWR programme. This did not deter the government. As Lawson² has explained: "The PWR was



Load duration curve for England and Wales for 1991-92 and 1995-96 (projected), showing the reduction in coal consumption by power stations because of the growth of gas-fired generation and imported electricity, and the protection of nuclear power.

seen as vital to demonstrate to the NUM that coal was not fundamental to the economy any longer, and that nuclear power stations could be built to time and to cost. The need for diversification of energy sources, the argument I used to justify the PWR programme, was code for freedom from NUM blackmail".

By the end of the Sizewell B public enquiry process in January 1987, the conditions necessary for the success of the PWR programme had gone. Changes in the British coal industry and the international coal market had greatly reduced the risk of NUM industrial action and seriously weakened the 'strategic diversity' argument for nuclear power. The lower price of coal, the availability of cheap gas for power generation, and the higher cost of capital invalidated the key economic assumptions behind the PWR programme. Already aroused by the Three Mile Island accident in the United States, public anxieties had been fanned by numerous press reports of radioactive emissions and links with leukaemia, and the fall-out from Chernobyl. Reprocessing and decommissioning costs were rising sharply and the original 15-GW programme had been reduced to a "small family" of four PWRs.

A worse blow was to come. In February 1988 the government adopted a unique structure for electricity privatization in England and Wales which was incompatible with including nuclear power in the sale and promoting its expansion. The new structure of the electricity supply industry ended the organizational and financial arrangements necessary to sustain nuclear power without overt protection and subsidy. The financial risks of existing nuclear stations and of investing in new PWRs were inimical to the shareholders' interest as there was now

no automatic 'pass-through' of costs to the consumer. Without guaranteed long-term power contracts, the private-sector owners could well fail to recover the capital cost of new nuclear stations (around £10-11 billion for the proposed four PWRs).

After the decision in November 1989 to keep nuclear power in the public sector, the government's justification for its continued support for nuclear power shifted away from 'low generating costs' to the need to retain the nuclear option to assist energy supply security and to combat the greenhouse effect. The emphasis changed from promoting nuclear expansion to defending the nuclear *status quo*, that is, the investments that had already been made. The defence strategy included an Exchequer (Treasury) provision of up to £2.5 billion towards back-end nuclear costs; the non-fossil fuel obligation which guarantees that all nuclear output is sold; the levy on fossil fuel generation which finances more than half of nuclear generating costs; the creation of two public nuclear corporations; reprocessing deals which underwrite the financial viability of British Nuclear Fuels Ltd (BNFL) and allow the extended operation of Magnox reactors; and the promise of an in-depth review in 1994 which would take account of the environmental and other 'externalities' claimed in its favour.

Pressing on with the PWR programme now presented major difficulties for the government due to the combination of low fossil-fuel prices, the "dash for gas" and the higher cost of capital. A decision to build further PWRs using a discount rate lower than in the private sector and/or incorporating fuel diversity and environmental 'credits' in the investment appraisal would be contested as further backdoor subsidies to nuclear power.

Building new PWRs without long-term power contracts (which the regional electricity companies would be unable to give) or without guaranteed subsidy until the capital cost is fully recovered (until 2010–15 for a PWR authorized in 1995) would further increase the Exchequer liability. The total liability for nuclear decommissioning costs now stands at £18–20 billion. (In approximate terms, the estimated decommissioning costs are £5.4 billion for the Magnox stations ($9 \times £600$ million), £4.2 billion for the AGRs ($7 \times £600$ million), £4.6 billion for Sellafield and other fuel-cycle facilities, and £4.5 billion for the UK Atomic Energy Authority and Ministry of Defence.) As owner of the nuclear corporations, the government is the ultimate bearer of this liability.

Extending the protection and levy to cover the life of new nuclear stations (to avoid increasing the Exchequer liability) would further antagonize major energy users, the generators and regional companies, and would be seen as a deliberate long-term market distortion. Substituting a carbon/energy tax for the levy might not be seen as a good idea unless the European Communities, the United States and Japan followed suit.

Disparity of treatment

Carrying out these policies for coal and nuclear power involved decidedly unequal treatment of the two industries, as pointed out on various occasions by the House of Commons Energy Committee. As a public corporation operating at arm's length with the nationalized generating boards, British Coal was subjected to strict financial disciplines and made to operate commercially. Until 1990, nuclear power was an internal part of the generating boards; its costs were hidden in their accounts and increases were passed automatically to the consumer as part of the overall cost of electricity generation. There were other disparities.

First, nuclear power investment was allowed a 5% discount rate for most of the period compared with 10% for coal-mining investment (and also for gas investment). New nuclear stations were uneconomic when the discount rate was raised to 8% in 1988 and would be more uneconomic in relation to the higher discount rates used in the private sector. Second, when fossil fuel prices fell in 1986, government support for nuclear power continued on the basis of a long-term view of its strategic benefits. The coal industry was required to adjust to international market forces, and no strategic benefit was recognized. Third, at electricity privatization, nuclear power was given a guaranteed market and financial assistance for 8 years. Coal was given only 3-year contracts with the generators, with declining real prices, no

direct subsidy and no protection against the "dash for gas". The nuclear 'subsidy' was agreed with the European Commission under a formula which allowed 20% of power station fuelling to be protected. In the United Kingdom this entitlement was used exclusively to support nuclear power. Germany used it to support coal.

Fourth, the government actively supported the case for proposed nuclear stations in the planning enquiry process and readily gave consents for Sizewell B and Hinkley Point C, when the economic case had been undermined by events. The two major deep-mined coal developments, Belvoir and Hawkhurst Moor, were both rejected on grounds of local environmental impacts. Fifth, there was huge disparity in the size and conditions attached to government-funded research and development in favour of nuclear power, even when efficient clean combustion became essential for future coal-fired generation and technical solutions were in sight. Finally, when environmental emissions became major international issues, the government emphasized the environmental case for nuclear power (particularly with respect to the greenhouse effect). It did nothing to mitigate the unfavourable image of coal even though coal was by no means the sole source, nor a growing source, of SO₂ and CO₂ emissions, and new clean-coal technologies had the potential to reduce these emissions significantly.

In the case of coal, government financial support facilitated contraction as opposed to maintaining output at a higher level. The grants were for the social costs of the closures. The write-offs of capital and past losses and of future liabilities were the recognition by the equity owner that by-gones are by-gones. In the case of nuclear power, there were some write-offs for the same reason, but the biggest factor was the recognition of the huge liabilities for back-end costs as a result of past decisions. The purpose was to minimize the effect on taxpayers by keeping the oldest nuclear stations in operation as long as possible on the basis of avoidable costs only, taxing consumers via the levy to contribute towards the unavoidable (fixed) costs, and postponing decommissioning far into the future so as to minimize the current cost. The levy and the accountancy for decommissioning were the ostensible means chosen to minimize the current costs of past mistakes and distribute them between consumers and taxpayers. Doubts have recently been raised, however, as to whether the levy was being used exclusively for back-end costs.

There was in both cases an element of true subsidy. The formal position was that each nationalized industry was required to show a minimum rate of return on its investment programme as a whole,

not on each individual project (which recognized that some investment is of a non-revenue earning kind). For most of the period this was 5%. We regard the use of a 5% discount rate for nuclear investment when competing industries were using much higher discount rates for revenue earning investment as tantamount to a subsidy which the government, as equity owner and banker, could have prevented had it wished. For coal there was some indirect subsidy via the joint understanding with the CEEB, and subsequently via the 3-year coal contracts imposed on the generators at privatization. These gave prices which, although reducing in real terms, were considerably above those of imported coal. The real extent of the subsidy is arguable, as a sudden large switch to imported coal, given the relatively narrow international steam coal market and restricted facilities at UK ports, might well have increased the delivered price of imported coal at UK power stations.

Results versus objectives

Until October 1992, the government's policy objectives for coal seemed to have been largely achieved. The share of domestically produced coal in the total fuelling of UK power stations by the mid-1990s would be around 25% compared with more than 80% in 1979–80 (see figure opposite). The power of the NUM had been broken. The "ultimate privatisation" seemed to be on the point of achievement. By contrast, all there was to show for the government support of nuclear power was one PWR nearing completion (Sizewell B) and one further site consent (for Hinkley C), but no new construction was possible before 1994 at the earliest.

There were various reasons for the marked difference in the achievement of the policy objectives. First, nuclear objectives proved inherently more difficult to achieve, partly because the 5 years in obtaining planning consent for Sizewell B delayed construction of the first PWR from a time of high fossil-fuel prices to one of low prices and higher costs of capital. Second, contraction of the coal industry went with the grain of market forces, but the economic justification of the proposed PWRs was progressively more difficult to sustain after the fall in fossil fuel prices and after the nuclear cost increases which were revealed during electricity privatization. Third, before October 1992, the political attraction of the coal policy objective was considerable and there was no political constraint, but nuclear power expansion was unpopular after the Three Mile Island accident in 1979 and especially after Chernobyl in 1986. Finally, the policy for coal was compatible with the government's policies of increasing

competition and of privatization, but the structure adopted for electricity privatization was incompatible with the policy of nuclear expansion unless special measures of protection were introduced.

Some explanations of why the government maintained its pro-nuclear policy after 1986, when there was no political or economic advantage in so doing, strike us as implausible. The 'strategic diversity' case was dubious before the miners' strike and largely void thereafter, as other more cost-effective routes to diversity were available and not pursued. The 'reduction of emissions' argument similarly was weak as more cost-effective remedies were available. The influence of the 'nuclear technocracy' is unlikely to have been a decisive factor, as it was not nearly as united and powerful as in the 1950s and 1960s. More plausible is the likelihood that Thatcher's predilection for nuclear power, which reflected her antipathy to Scargill and the nationalized coal industry, had a decisive influence on her cabinet. It is also possible that government ministers genuinely believed that important countries should have nuclear power programmes, that nuclear power would provide the cheapest electricity in the long run, or (more recently) that preserving the nuclear option was important for environmental reasons.

There has been another powerful and continuing pressure on the government to continue to support the nuclear industry. Not only does a government so long in office bear direct responsibility for the nuclear corporations it created or restructured (Nuclear Electric, Scottish Nuclear, BNFL, the NNC, the UK Atomic Energy Authority and NIREX), but it is also faced with the huge back-end liabilities which would fall on the Treasury if the nuclear industry were rapidly run down. Funding the provisions necessary to meet the decommissioning liabilities of £18–20 billion, even under the present policy of postponing decommissioning for more than 100 years for reactors and 50 years for the reprocessing plant at Sellafield, would substantially increase public expenditure at a time the large public sector borrowing requirement is of acute concern to the government and the City of London. The government therefore has a strong interest in maintaining existing nuclear capacity in a way that reduces or postpones the net liabilities. This includes continuing the operation of the Magnox and AGR stations as long as compatible with safety, and underpinning the new reprocessing plant (Thorp) with long-term reprocessing contracts with captive home customers despite the availability of the allegedly cheaper dry storage option. The public sector borrowing requirement argument is also consistent

Symbol of a bygone era? A pithead stands stark against a Yorkshire dawn.

with the changes taking place in the nature of the government support. Previously, 'support' meant promotion of a programme of new stations, financed through the public sector and backed up by very large public-sector funded research and development. 'Support' now involves financial and other arrangements to safeguard the continued operation of existing nuclear stations, to postpone liabilities as long as possible, and to postpone further commitments.

National Interest assessment

The threat posed by the NUM has gone. The costs of fuel for power generation are now considerably lower in real terms than in 1979–80, and the electricity industry has a more diverse mix of fuel supplies and sources. But little of this results from government policy on nuclear power since 1979. Lower fuel prices stem mainly from changes in the international fuel market since the mid-1980s. The diversification of power-station fuel supplies at the end of the period was not due to nuclear expansion but to other factors, particularly electricity privatization and the earlier decision to install a transmission link to France for load-balancing purposes.

The negative effects of government policies for coal and nuclear power are more evident than the positive effects. In the absence of major new policy initiatives, the deep-mined coal industry would run down to 30 million tonnes or less in the mid-1990s and further still by 2000. There would be little chance of recovery as that would require huge long-term investment to create major replacement capacity. Such investment is unlikely without long-term coal contracts, especially by more risk-averse private owners. The closures announced on 13 October 1992 would preclude giving the 31 pits time to continue reduc-

ing costs to become competitive with imported coal. The "dash for gas" seems likely to result in colliery closures excessive in relation to coal production costs and the costs of the unemployment involved. British Coal's expertise in the development of clean coal technology is on the point of being lost.

Ironically, while the government championed market forces, the one long-term element of government policy — maintaining nuclear power capability — has required arrangements 'outside the market'. However, the nuclear industry faces many uncertainties. The breakthrough with nuclear capital costs has not been achieved, and it will not be known until Sizewell B has been operating for some years whether the PWR will become the established nuclear option that has been sought for 30 years. There is uncertainty over the framework of financial support and over the future of reprocessing, and therefore of BNFL, and the financial risks of new stations.

It is not easy to see how the manifest disparity of policy towards the coal and nuclear industries could be justified, particularly in the period after the 1984–85 NUM strike. The government has not explained what compelling strategic or economic reasons there were to favour nuclear power rather than coal. Explaining it all in terms of Scargill may have had some logic before the strike of 1984–85. Afterwards it did not. □

Mike Parker and John Surrey are at the Science Policy Research Unit, University of Sussex, Falmer, Brighton BN1 9RF, UK. This article is based on Unequal Treatment: British Policies for Coal and Nuclear Power 1979–92 SPRU, 1992, price £14.

1. *Nature* **361**, 283–284 (1993).
2. Lawson, N. *The View from No. 11* (Bantam, London, 1992).
3. Parkinson, C. *Right at the Centre* (Weidenfeld and Nicholson, London, 1992).

Pondering on British science policy

A brief report of last week's meeting on the future of British science administration reveals differences of opinion that will not be easily bridged.

Edinburgh. The British government, headed at the time by Prime Minister Margaret Thatcher, ignored the warnings of its top scientific advisory committee when it decided in 1987 to withdraw its support for 'near market' research. Six years later, the need to remedy the subsequent neglect of applied research has become a central focus for the government as it draws up its promised White Paper (policy document) on the organization of science.

A succession of speakers at a meeting organized in Edinburgh last week by the Edinburgh International Science Festival and *Nature* urged the government to turn its back on Thatcher's approach and take firmer steps to stimulate the links between industry and the academic world. One of these was Sir Graham Hills, vice-chancellor of the University of Strathclyde until two years ago and now a member of body whose earlier advice was ignored, the Advisory Council on Science and Technology (ACOST).

Referring to the 1987 decision — which had, for example, led to the virtual drying up of research council funds for applied research in universities — Hills said that "the one thing that the UK was poor at was near market research". ACOST had tried to block the government's move; but "it was a battle fought and lost", he said.

Three other needs emerged as key points that speakers at the meeting urged the government to address in the White Paper: an improved career structure for scientists, particularly for postgraduate students employed on short-term contracts; an assurance that any reorganization of the research councils will protect the science base and not become too "technology driven"; and closer coherence between national and European science policy.

Hills defended proposals made by ACOST, in its submission to the government on the White Paper, to separate funding bodies such as research councils from the institutions performing research. Such a separation, he suggested, would sharpen the value judgements made by scientists in seeking funds. Furthermore, a previous concern with the inputs into scientific research had resulted in a relative neglect of managing the outputs, he said; this had impaired the relationship between industry and government as 'customers' of research, and the scientific community as 'providers'.

Hills described ACOST's proposals to reorganize the five current research councils into two new bodies — the Council for the Advancement of Scientific Knowledge and

the Council for Mission Oriented Research — as a first attempt to "bring respect" into supporting and managing research linked to the needs of industry by clarifying the purpose of research and distinguishing different funding routes.

But this approach was criticized by David Wallace, professor of physics at the University of Edinburgh, and chairman of the science board of the Science and Engineering Research Council (SERC). Wallace said that separating funding for basic and applied research would have two major defects. First, it implied a single research council covering all fields of basic science; "I do not believe that is practical".

Second, it was based on the false premise that there is a linear development leading from basic research through strategic and applied research to the development of products. But the SERC's experience, which showed areas in which basic research (for example in mathematics) could sometimes have immediate applications, showed that "it does not always work like that", said Wallace. He said industrial collaboration could be dangerous if it led to a particular research project being "technology driven".

There was general agreement at the meeting that a new initiative is needed from the government to provide greater infrastructural support for linking basic and applied research, as the Fraunhofer Institutes do in Germany (and as the proposed Faraday Centres, which now appear to have been abandoned, would have done in Britain). Several participants praised the LINK scheme, run with support from the Department of Trade and Industry, as a step in this direction, despite its excessive bureaucracy.

But there was also a strong feeling that such initiatives should be built on the resources and capabilities of universities. "Contact between applied research and basic research raises the standards of the former and introduces realistic priorities into the latter" said Max Irvine, principal of the University of Aberdeen. "Universities would seem to be an ideal setting for such contacts to flourish."

Irvine said he applauded signs of a greater public commitment to near market research, describing the dominance of the Stock Market as the "major deterrent" to industrial investment in research, since, in contrast to Japan, it demanded too high a return. But resources to provide this commitment should not be taken entirely from within the public sector science vote; funds should be transferred from other government departments if the country's capacity for basic science "is

not to be crippled".

A strong plea for better treatment of postgraduate students was made by David Glover, head of the Cancer Research Campaign's cell cycle group at the University of Dundee. "The lack of career structure is evident from research student days onward", said Glover. One problem was the wide disparity in student stipends. He urged the government to follow the example given by medical charities, whose current level of stipends for postgraduate students is about twice that offered by the research councils. He added that the career structure of scientists in universities and research institutes needed to be carefully re-examined if Britain was to make best use of its limited resources of trained personnel.

Another area needing attention, said Glover, was the interaction between Britain's national research strategy and that originating from the Commission of the European Communities (EC) in Brussels. An increasing proportion of Britain's research resources were being directed to EC projects that had a short life span and were often poorly thought out, he said. "The programmes are highly bureaucratic, and neglect career structure even more than our national programmes."

Further support for closer attention to this area came from Wallace, who questioned the way in which William Waldegrave, the cabinet minister responsible for science, had asked for comments in preparing his White Paper on how British research could best "complement" that conducted through European programmes. "What we should really be seeking is a coincidence of purpose, rather than complementarity", he said.

Speaking from the floor, Joe Lamb, professor of physiology at the University of St Andrews, and chairman of the Save British Science campaign, said that one important problem that needed to be tackled was "a lack of vision" in most sectors of industry. This was despite the fact that, if analysed on a sector-by-sector basis, those companies which had the highest expenditure on research and development — in particular in the chemicals and pharmaceuticals industry — were also the companies that were the most successful in international markets.

The message that left the meeting was that the Thatcher 'mini-revolution' in leaving the stimulation of technology (and thus indirectly of research) to the market-place has failed. New government measures are needed, and hopes are high that Waldegrave is prepared to deliver.

David Dickson

Penetrating the peroxisome

David Valle and Jutta Gärtner

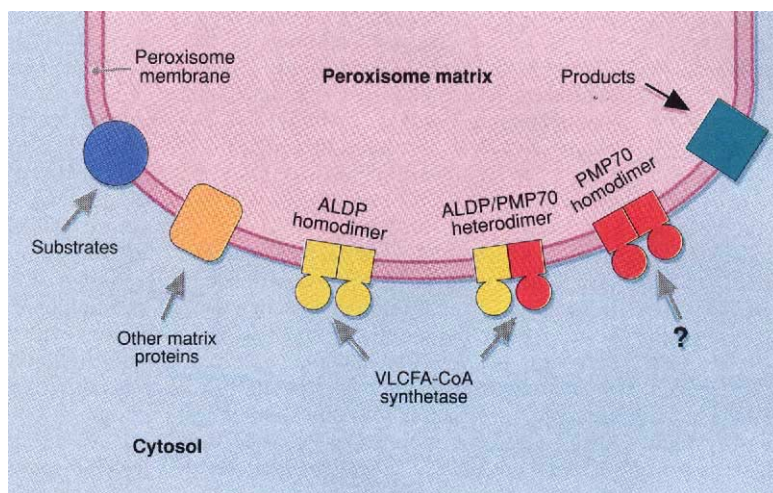
In the past twelve months, the molecular disciples of Sir Archibald Garrod¹ have invaded the world of peroxisomes. Three genes involved in inborn errors of peroxisomal function have been identified²⁻⁴, the most recent and surprising example being reported on page 726 of this issue⁴ by Mosser *et al.* The disorder concerned is X-linked adrenoleukodystrophy.

Peroxisomes, or microbodies as they were often called, are single, membrane-bound organelles present in nearly all eukaryotic cells. They were recognized as biochemical and morphological entities only 30 years ago in the pioneering studies of De Duve and his colleagues⁵. In mammalian cells their size (0.1 to 1.5 μm) and number vary depending on the type of cell and its physiological state; they are particularly abundant in the liver and kidney. The peroxisome membrane contains two quantitatively major (M_r 22K and 70K) and several less abundant peroxisome-specific proteins. The organelle matrix has more than 40 enzymes which catalyse a variety of reactions⁶, both anabolic (plasmalogen biosynthesis, bile-acid biosynthesis) and catabolic (β -oxidation of very-long-chain fatty acids, H_2O_2 -based respiration). The metabolic interaction of these intraperoxisomal reactions with extraperoxisomal pathways is extensive, necessitating transport of a variety of metabolites across the peroxisome membrane (see figure).

Interest in peroxisomes stems not only from their biology, but also because in several genetic diseases there is either isolated deficiency of a specific matrix enzyme (single-enzyme deficiencies) or defects in the formation of the organelle with deficiency of several peroxisomal functions (inborn errors of peroxisome biogenesis). X-linked adrenoleukodystrophy (ALD) has been the exemplar for the former, and Zellweger syndrome for the latter^{7,8}.

ALD is the most common inherited peroxisomal disorder. Patients with the severe phenotype, the childhood cerebral form, seem unaffected until mid-childhood, when there is onset of adrenal insufficiency and progressive

neurological dysfunction, leading to death within a few years. Interestingly, a much milder phenotype with onset in adult life, adrenomyeloneuropathy, segregates in some of the same families. In both forms, β -oxidation of very-long-chain fatty acids (VLCFA), particularly hexacosanoic acid (C26:0), is impaired. In studies with patients' fibroblasts, the biochemical defect was localized to the level of the peroxisomal lignocero-



Hypothetical model of peroxisomal membrane transporters. ALDP may assemble into a functional transporter either by homodimerizing or by heterodimerizing with PMP70 or some other ABC 'half transporter'. VLCFA-CoA synthetase is the expected ligand. Functional and genetic studies suggest that there are several other peroxisomal membrane transporters for matrix proteins and for the products and substrates of peroxisomal matrix enzymes. ALDP, adrenoleukodystrophy protein; PMP, peroxisomal membrane protein; VLCFA, very-long-chain fatty acids.

CoA synthetase (also known as VLCFA-CoA synthetase⁶).

As early as 1981, Migeon and colleagues found linkage of ALD to markers at position 28 on the long arm of the X chromosome (Xq28)⁹. Purification of VLCFA-CoA synthetase proved difficult, so Mosser *et al.*⁴ used the positional cloning approach. Surprisingly, the gene they found encodes a protein (ALDP) with homology to a 70K peroxisomal membrane protein (PMP70), rather than to the expected VLCFA-CoA synthetase. The genetic evidence that this gene is responsible for ALD is convincing; six of 85 ALD probands have intragenic deletions.

Both PMP70 and the newly discovered ALDP belong to the ATP-binding cassette (ABC) transporter protein superfamily¹⁰. Members of this family are membrane proteins involved in transport of a wide variety of ligands ranging in size from ions to proteins in excess of 100K. The typical structural motif of an ABC transporter consists of a highly hydrophobic amino-terminal

domain with six putative membrane-spanning segments, followed by a hydrophilic domain containing an ATP-binding fold or cassette. In many ABC transporters this motif is repeated in tandem so that a single polypeptide has four domains, two hydrophobic and two hydrophilic (the cystic fibrosis transmembrane conductance regulator protein and multiple-drug-resistance protein are well-known examples of this sort). By contrast, in size and organization ALDP and PMP70 are 'half transporter' versions with one hydrophobic and one hydrophilic domain.

The similarity in sequence between ALDP and PMP70 is impressive; overall, they have 38.5 per cent of amino acids in common, which increases to 56 per cent in the 210 residues making up their carboxy-terminal hydrophilic domains. ALDP is more akin to PMP70 than to other members of the ABC transporter superfamily. This high degree of relatedness, the known peroxisomal membrane location of PMP70, and the nature of the biochemical abnormality in ALD, make it tempting to conclude that ALDP is a peroxisomal membrane protein. But this conclusion is at odds with inferences from the northern blot data provided by

Mosser *et al.*⁴, which show that some tissues with abundant peroxisomes (liver and kidney) have low levels of the ALDP transcript, whereas other tissues with fewer peroxisomes (cardiac and skeletal muscle) have high levels of the ALDP transcript. Obviously, the question of subcellular localization of ALDP should be addressed directly in short order.

This caveat aside, the hypothesis that ALDP is a peroxisomal membrane transporter directly involved in the import of VLCFA-CoA synthetase is appealing. Furthermore, the extent of the sequence similarity between ALDP and PMP70 raises the possibility that they might interact directly to form a heterodimeric transporter. Four domains (two hydrophobic and two hydrophilic) are thought to be a requirement for ABC transporter function¹⁰. So, dimerization of some sort is likely for both ALDP and PMP70. Two other 'half transporter' ABC proteins, TAP1 and TAP2, with a degree of relatedness similar to that of ALDP to PMP70, heterodimerize to form a trans-

porter which imports antigenic peptides from the cytosol to the lumen of the endoplasmic reticulum^{11,12}.

If ALDP and PMP70 do indeed heterodimerize to form the VLCFA-CoA synthetase transporter, additional questions must be answered. Why do mutations in the former produce ALD, whereas mutations in the latter appear to result in Zellweger syndrome, an autosomal inborn error of peroxisome biogenesis, lethal in infancy? Perhaps PMP70 also dimerizes with itself or different partners in the peroxisome membrane to form transporters required for import of other peroxisomal enzymes. If this were the case, deficiency of PMP70 would have a more profound effect on assembly of the organelle and, consequently, on the organism. The patient's death would precede development of symptoms due to impaired VLCFA oxidation.

As is often the case with such developments, the discovery of the ALDP gene opens another chapter in the research agenda. Is ALDP really present in the peroxisomal membrane, and how and in what form does it function? Does it homodimerize or does it have a partner protein and, if so, is that partner PMP70? If a defect in the transport of VLCFA-CoA synthetase into peroxisomes results in deficiency of enzyme activity and hence the ALD phenotype, where are the patients with mutations in the structural gene for this enzyme? Lastly, will improved understanding of the pathophysiology enable us to develop an effective treatment (which is not yet available, for all the excitement surrounding the film *Lorenzo's Oil* — see the review on page 695 of this issue)? The fact that ALD patients are well for the first few years of their life offers a window of opportunity for instituting therapy. What we have to do is to find that therapy. □

David Valle and Jutta Gärtner are in the Department of Pediatrics and the Howard Hughes Medical Institute, Johns Hopkins Medical School, Baltimore, Maryland 21205, USA.

Natural selection in the corona

Kenneth H. Schatten

"HEAT always passes from the hotter to the cooler" according to the old song¹ — so why is the Sun's corona hotter, by several million degrees, than the underlying Sun, where all the energy is generated? Jack Scudder has a new suggestion². Where others have sought to put extra energy into the corona through mechanical or magnetohydrodynamical means, Scudder suggests that

like the Earth's atmosphere, be very thin, owing to the enormous solar gravity. It should not extend as far as the observations indicated (many solar radii). A number of possibilities emerged: that the light was scattered not from the solar atmosphere, but rather from the Moon's atmosphere; that the gases observed were of a new material, 'coronium'; or later, that the Sun's atmosphere was extremely hot (several million kelvin).

With the advent of the space age, outflowing gases were observed to form a solar wind, previously predicted by Parker³. Parker's theory was based upon the supposition of a hot corona, and was supported by Biermann's observations of apparently wind-swept comet tails.

So the hot corona explained the extension of the Sun's atmosphere and the solar wind, but it created a new problem: how could the Sun's atmosphere, heated solely by the Sun, be hotter than the Sun's surface, at 5,800 K? Thermodynamics tells us that, within its domain of applicability (local thermodynamic equilibrium), one cannot transfer heat from

one location to another at a higher temperature, without doing work. It turns out that the need for equilibrium is the chink in the armour of thermodynamics that allows Scudder's model to succeed where others failed. Without thermodynamic equilibrium, temperature itself is not well defined.

For many years, attempts to uncover

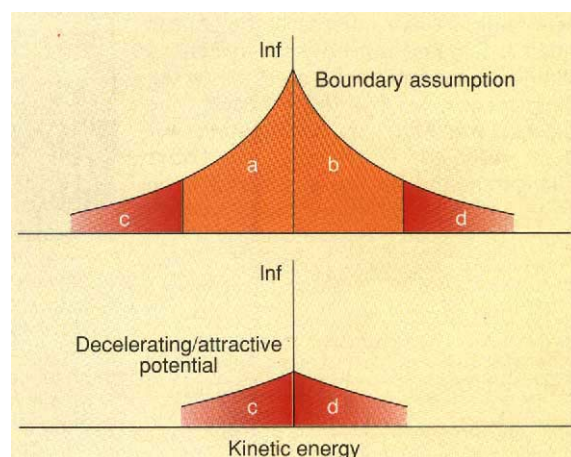


FIG. 1 With a total eclipse blotting out the bright disk of the Sun's photosphere, the hot, tenuous corona can be seen. Scudder argues that all stars should have such a corona, explaining their X-ray emissions.

a velocity filter is at work — allowing the fastest, and so most energetic and hottest, ions to rise into the Sun's atmosphere.

The high coronal temperature is a mystery wrapped in an enigma; the problem arose when it was first realized that the atmosphere of the Sun, seen during solar eclipses (Fig. 1), should,

FIG. 2 The kind of non-equilibrium velocity distribution that might give a hot corona by Scudder's velocity-filtration mechanism. Top, the velocity distribution in the photosphere; below, the deceleration and filtered component (c,d) that reaches the corona; the low-energy core (a and b) is removed. In this logarithmic plot, an equilibrium, maxwellian distribution would be strictly triangular, with constant slopes. In that case, filtration would give the same distribution, with no change in temperature.



1. Scriver, C. R. & Childs, B. *Garrod's Inborn Factors in Disease* (Oxford University Press, 1989).
2. Shimozawa, N. *et al. Science* **255**, 1132–1134 (1992).
3. Gärtner, J., Moser, H. & Valle, D. *Nature Genet.* **1**, 16–23 (1992).
4. Mosser, J. *et al. Nature* **361**, 726–730 (1993).
5. Baudhuin, P. *Physiol. Rev.* **46**, 323–357 (1966).
6. van den Bosch, H. *et al. A. Rev. Biochem.* **61**, 157–197 (1992).
7. Moser, H. W. & Moser, A. B. in *The Metabolic Basis of Inherited Disease* (eds Scriver, C. R. *et al.*) 1511–1533 (McGraw-Hill, New York, 1989).
8. Lazarow, P. B. & Moser, H. W. in *The Metabolic Basis of Inherited Disease* (eds Scriver, C. R. *et al.*) 1479–1509 (McGraw-Hill, New York, 1989).
9. Migeon, B. R. *et al. Proc. natn. Acad. Sci. U.S.A.* **72**, 5066–5070 (1981).
10. Higgins, C. F. *A. Rev. Cell Biol.* **8**, 67–113 (1992).
11. Kelly, A. *et al. Nature* **355**, 641–644 (1992).
12. Spies, T. *et al. Nature* **355**, 644–646 (1992).

the source of the high coronal temperature relied on thermodynamics. Researchers examined various mechanisms (plasma instabilities, wave heating; nanoflares, shocks, and so on) to convert the Sun's turbulent energy in the convection zone into heat in the distended corona. The stumbling block was finding a mechanism that could transport the mechanical energy up from its source, thermalize it into heat in the corona, and yet sustain a sharp temperature transition as observed.

Scudder has upset the applecart by forsaking the transport/thermalize approach. He examines the consequences of a nonthermal velocity distribution that might exist near the base of the transition region (top of the chromosphere) as a by-product of the incomplete dissipation of power in the convection zone beneath. The main temperature inversion effect may be understood simply as follows. Near the Sun's surface, gravity pulls the turbulent gases inward. Particles with the lowest energies will have the greatest difficulty in escaping. With increasing altitude, this process smoothly removes or filters out the lowest part of the random distribution of speeds, while degrading the kinetic energy of those that can rise against the pull of gravity. This process leads to a rise in the effective temperature of the gases (Fig. 2). The mechanism works only if there is a nonthermal velocity distribution in the photosphere. With the Sun's photosphere in thermodynamic equilibrium, removing the core of the velocity distribution would not affect the temperature of the remaining velocity distribution. The nonthermal wings of the velocity distribution are vital to the model. Scudder skirts the domains of thermodynamics' applicability by allowing only high-quality energy (the hot gases) into the upper solar atmosphere.

This mechanism also explains the sharp nature of the temperature inversion above the Sun's transition region. Scudder finds that the temperature can rise abruptly from the temperature minimum, near 4,500 K, to values in excess of 1 million kelvin over scale heights as small as 200 km. Indeed, such inversions should be formed above all stars on the main sequence by this process, he argues, providing an explanation for their observed X-ray emissions. Other nuances of the heating problem — the large ultraviolet Doppler width of spec-

tral lines, a correlation of temperature excess with magnetic field orientation, the non-maxwellian tail of the velocity distribution, a proportionality between mass and temperature of minor ions in the solar wind — all fall into place. Seen individually, they have been disjointed facts. Of course, a degree of scepticism is always in order, and the work must be checked and tested, but it is very persuasive.

Why did Scudder find his solution when others failed? In fact, others have examined this approach⁴⁻⁶, but lacked a particular insight. Scudder had a unique perspective on the problem. His previous experience with NASA's space

experiments in the solar wind, involving collisionless shocks⁷, had shown how non-equilibrium distributions can develop. Ascertaining how a plasma really works in space (how shocks behave), rather than on the ground where the large-scale collisionless plasma can never be viewed, appears to have paid off. With this experience, Scudder, like an alert quarter back, has made a lightning 'end run' (to borrow the term from American football), bypassing the thermodynamic defence line. □

Kenneth H. Schatten is at the Laboratory for Planetary Atmospheres NASA/GSFC, Greenbelt, Maryland 20771, USA.

DNA REPLICATION

Methyltransferases in foci

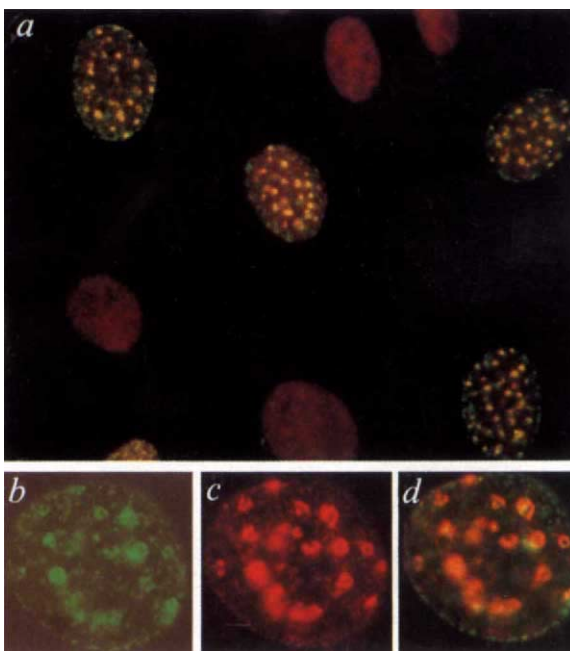
J. Julian Blow

THE many thousands of replication forks involved in duplicating large eukaryotic genomes are not distributed randomly throughout the cell nucleus but are clustered together in some 50–250 foci¹. Factors directly involved in DNA replication²⁻⁴ are found within these foci. What, however, might a protein not so involved be doing there? This is the question raised by observations pub-

lished in *Cell*⁵, which in turn bears upon the way DNA replication is co-ordinated in eukaryotic cells.

The protein concerned is DNA methyltransferase (MTase), an enzyme that methylates cytosine residues in DNA to form 5-methylcytosine. Most methylcytosine residues occur in the palindrome 5'-C-p-G-3', with methylation on both strands. Not all CpG palindromes are methylated, and the methylation state of each CpG is potentially heritable: after replication, both daughter strands are hemimethylated, and the full-length MTase protein can re-methylate them to restore the parental pattern.

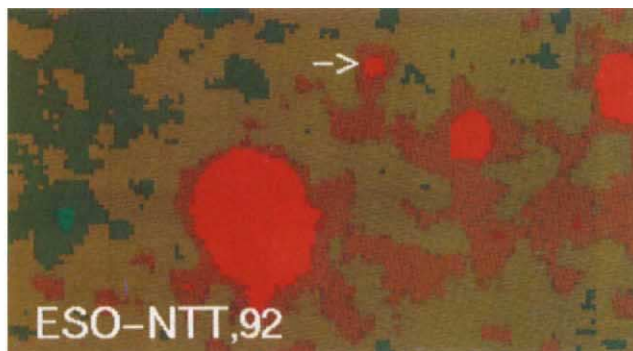
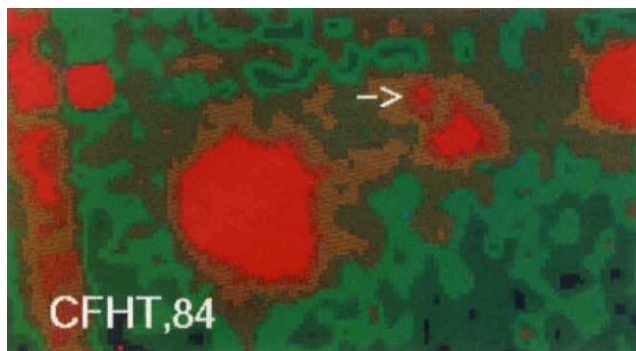
The exact function of methylation in eukaryotic cells is not known, although most of the evidence points to a role in transcriptional control. Many lower eukaryotes appear to contain no methylcytosine at all. Further, a mouse cell line homozygous for a partially inactivating deletion within the MTase gene seems to be quite unaffected by this defect⁶. But when the cell line was used to create chimaeric mice, the homozygous MTase deletion was lethal for embryos. This result is consistent with an inability of mutant embryos to regulate transcription at stages when differential transcription becomes crucial for



Co-localization of MTase (red) and sites of DNA synthesis (green) in mouse 3T3 cells. *a*, Survey micrograph of log-phase cells pulsed for 15 minutes with the nucleotide precursor BrdU. 70–80 per cent of the cells are in G1 and G2 phases of the cell cycle and so do not incorporate BrdU. These cells show diffuse MTase staining. The 20–30 per cent of cells in S phase show marked co-localization of MTase and BrdU. *b–d*, High-power view of a single S-phase nucleus showing BrdU (*b*), MTase (*c*) or BrdU plus MTase (*d*). (Reproduced from ref. 5.)

1. Flanders, D. & Swan, M. *At the Drop of Another Hat* (Parlophone, London, 1964).
2. Scudder, J. D. *Astrophys. J.* **398**, 299–318; 319–349 (1992).
3. Parker, E. N. *Astrophys. J.* **128**, 664–674 (1958).
4. Brandt, J. & Cassinelli, J. P. *Icarus* **5**, 47–63 (1966).
5. Lemaire, J. & Scherer, M. *Rev. Geophys. Space Phys.* **11**, 427–468 (1973).
6. Schatten, K. H. & Mayr, H. G. *Astrophys. J.* **283**, 853 (1984).
7. Scudder, J. D. *et al. J. geophys. Res.* **91**, 11019–11052 (1986).

Catch a moving star



WHEN it became clear last year that the mysterious γ -ray source Geminga is a γ -ray pulsar at an apparent distance of less than 400 parsecs, the prospects for measuring its proper motion (perpendicular to the line of sight) looked promising. On page 704, Bignami *et al.* show that the motion of the putative optical counterpart to Geminga, G'', can be seen by comparing images constructed from data from 1984, 1987 and 1992 (the first and

last of which are shown above). The magnitude of the proper motion is consistent with the idea that G'' is a neutron star (making the association with Geminga likely) about 100 parsecs away. Extrapolation 300,000 years back to the object's position when it was formed in a supernova suggests that this outburst may have blown the bubble in the interstellar medium whose origin has been a matter for debate (page 706). P.B.

embryonic development, an interpretation that underscores the surprising nature of the localization of MTase at replication foci. Indeed, why should replication forks be clustered into such foci in the first place?

Part *a* in the figure, taken from the paper by Leonhardt *et al.*⁵, shows cells double-labelled with antibodies against the MTase protein (red) and newly synthesized DNA (green). Cells in the G1 or G2 phases of the cell cycle show a diffuse nuclear MTase staining, but S-phase cells, undergoing DNA replication, give a pronounced spotty pattern for both MTase and newly synthesized DNA. The coincidence of the two labels produces a yellow colour. Part *b* shows sites of DNA replication and part *c* shows MTase distribution in a single late S-phase nucleus. The two labels are highly co-localized, as shown by part *d*, which is a superposition of *b* and *c*. While the large toroidal replication foci, which contain centromeric heterochromatin, clearly contain MTase protein, the smaller peripheral foci also stain weakly for MTase. Whether MTase associates with all replication foci to a similar degree, or whether it has a preference for centromeric chromatin is unclear, though co-localization of MTase and replication was observed throughout S phase⁵.

Leonhardt *et al.*⁵ find that the ability of MTase to localize to replication foci is conferred by a 250-amino-acid signal sequence at its amino terminus. Targeting does not appear to be required for enzymatic activity, because certain deletions are not targeted (they show a diffuse nuclear staining throughout S phase) but retain full activity with substrates *in vitro*. This suggests that MTase localization to replication foci is not caused simply by the abundance of hemi-

methylated substrate at these sites, but is produced for a specific reason. One likely function that targeting might achieve is to increase the rate of methylation of hemi-methylated sites generated by the replication forks. Increased efficiency for replication-fork proteins by physical association was first suggested for the bacteriophage T4 replication fork, whose component proteins associate into a miniature 'protein machine'⁷. But although efficient use of DNA polymerases at a replication fork is required for rapid movement of the fork, it is unlikely that rapid methylation would speed up the process of replication. Further, this does not explain why replication forks (and associated MTases) should be clustered together within the nucleus.

Duplication of the eukaryotic chromosome during the cell cycle requires not only the replication of DNA but the duplication of a number of other features such as DNA methylation, chromatin modification and the presence or absence of transcription factors. Duplication of these chromosomal features is likely to be a sequential process, which might be the key to understanding why MTase is so closely associated with replication foci. Methylation of DNA occurs more rapidly in S-phase cells, possibly because of the relative ease of access of MTase to the DNA when other chromatin proteins are no longer tightly associated with it⁸. Certain proteins can then bind specifically to DNA containing methylated CpG, and this binding seems to be responsible for the repression of transcription that occurs at methylated promoters^{9,10}. The presence of such factors is also likely to mediate the degree of overall compaction that the replicated chromatin subsequently undergoes.

The organization of replication forks

into clusters may serve to concentrate these sequential processes into functional zones, for instance by allowing DNA methylation to occur before binding of proteins to the nascent DNA. Creation of such zones could also serve to exclude unwanted activities from the region of DNA polymerization; for example, the gapped and unwound DNA structures present at the replication fork might elicit an inappropriate response if detected by DNA repair enzymes or by DNA-binding proteins such as poly-(ADP-ribose) polymerase¹¹.

The close association of DNA methyltransferase with replication foci emphasizes that the process of chromosome replication in eukaryotic cells is more than the simple duplication of the two strands of DNA, but involves the duplication of other epigenetic features as well. It seems increasingly likely that the spatial organization of replication forks into replication foci plays a part in the co-ordination and control of this sophisticated process. □

J. Julian Blow is a Lister Institute Fellow at the ICRF Clare Hall Laboratories, South Mimms, Potters Bar, Hertfordshire EN6 3LD, UK.

1. Nakamura, H., Morita, T. & Sato, C. *Exp. Cell Res.* **151**, 123–133 (1986).
2. Bravo, R. & Macdonald-Bravo, H. *J. Cell Biol.* **105**, 1549–1554 (1987).
3. Wilcock, D. & Lane, D. P. *Nature* **349**, 429–431 (1991).
4. Adachi, Y. & Laemmli, U. K. *J. Cell Biol.* **119**, 1–15 (1992).
5. Leonhardt, H., Page, A. W., Weier, H. & Bestor, T. H. *Cell* **71**, 865–873 (1992).
6. Li, E., Bestor, T. H. & Jaenisch, R. *Cell* **69**, 915 (1992).
7. Alberts, B. M. *Cold Spring Harbor Symp. quant. Biol.* **49**, 1–12 (1984).
8. Kastan, M. B., Gowans, B. J. & Lieberman, M. W. *Cell* **30**, 509–516 (1982).
9. Boyes, J. & Bird, A. *Cell* **64**, 1123–1134 (1991).
10. Levine, A., Cantoni, G. L. & Razin, A. *Proc. natn. Acad. Sci. U.S.A.* **88**, 6515–6518 (1991).
11. Satoh, M. S. & Lindahl, T. *Nature* **356**, 356–358 (1992).

Predicting large tsunamis

Emile A. Okal

In 1992, two catastrophic tsunamis killed more than 2,000 people. These giant sea waves, generated by earthquakes or submarine landslides, can travel across the oceans and bring destruction to faraway shores: in 1960, an earthquake in Chile led to the deaths of more than 100 people in Japan. In the case of the Nicaraguan earthquake on 2 September last year, the death toll was high because the ground tremors felt by the coastal populations were deceptively mild. On page 714 of this issue¹, Kanamori and Kikuchi take a hard look at the seismic data available for this earthquake. The long-period seismic waves measured at the time should, they say, have conveyed the much-needed warning.

In the Nicaraguan event, there was an alarming disparity between the fairly low conventional magnitude (reported to be as low as $m_b = 5.3$) and the size of the tsunami generated: in some sections of the coast, the earthquake was not even felt by the local people who were swept away by the waves a few minutes later. The other major tsunami of 1992, on 12 December, completely destroyed the village of Riangkrok on Flores Island, Indonesia (see the accompanying box). The waves at some locations reached 26 metres in height, whereas along most of the shoreline the maximum amplitudes were only 3 to 4 metres.

Because tsunami waves cross the ocean comparatively slowly (250 m s^{-1} , or roughly the speed of a jetliner), early warning should be possible. The risk could be assessed either from reports of the size of sea waves at shorelines closer to the epicentre, or by evaluating the parent earthquake, whose surface waves travel about 15 times faster. Not surprisingly, tsunami generation is often directly related to the earthquake's size — in layman's terms, its magnitude. But the concept of magnitude was developed at a time when there was little theoretical understanding of the propagation and generation of seismic waves, or of the physical nature of the forces required to describe an earthquake source. The chief shortcoming of the conventional 'Richter' magnitude scale, traditionally based on the magnitude at a period of 20 seconds, is that it saturates around $M = 8$, just where tsunami generation can become substantial. (The body wave magnitude, m_b , is measured at a period of about 1 second, and saturates still earlier.) Far better as a measure of earthquake size is the seismic moment, a *bona fide* physical quantity associated with the earthquake source, and measured in physical units (dyn cm). This can

be used to extend the concept of magnitude to much lower wave frequencies, thus avoiding the problem of saturation.

Thanks to developments in broadband seismic instrumentation in the 1980s, we can now estimate the seismic moment of a parent earthquake in quasi-real time. Distant coastal populations can then be warned if tsunamis seem likely. Several algorithms exist²⁻⁴, one of which has been successfully tested down to a distance of 1.5° (about 170 km)⁵. If the quake is closer than this, however, there may be only a few minutes in which to assess the earthquake and deliver a warning. Many coastal communities in earthquake-prone areas rely on public education, which can be summarized as "If it shakes really strongly, run

for the hills at once".

Enter the ominous 'tsunami earthquakes'. This term was coined by Kanamori in the 1970s to describe events whose tsunami was much larger than expected from their seismic waves⁶. Examples include the 1896 earthquake in Sanriku, Japan; the 1946 Aleutian earthquake (which unleashed possibly the strongest tsunami of the century in the Pacific, despite a Richter magnitude of only 7.2), and smaller events in the Kuriles (1963, 1975) and Peru (1960).

Part of the problem with tsunami earthquakes is that their size may be underestimated if the seismic energy is released exceedingly slowly. A slow source is inefficient at generating the high frequencies that rock buildings and alert their inhabitants (0.2 Hz and above) and even those used in conventional magnitude scales (0.05 Hz and above). But the vibrations can still interfere constructively at the mantle-wave

Disaster on Flores Island



THREE weeks after the tsunami hit Flores Island, Indonesia (see Okal's article above), an international tsunami survey team of scientists and engineers from five countries visited the location. They found that, at the site of the village of Riangkrok, located on the northern tip of the easternmost peninsula, the waves had caused devastation at enormous heights above sea level. The maximum height reached was measured as 26.0 m on the south hillside slope; the average of four different measurements was 19.6 m from sea level at the time of the tsunami onslaught. There is now no trace of the village at Riangkrok. It was destroyed and completely washed away. Almost all the coconut trees have been knocked down and washed away, leaving only their root marks still visible on the beach. The area engulfed by the tsunami is unmistakably identifiable as brown, bare ground. The size of the debris and coral rocks (dragged up from the sea floor) can be seen by comparison with the person to the right of the rock in the picture. This sort of destruction is very different from that observed at other tsunami-damaged locations in Flores Island, where the waves ran up to heights of only 3 to 4 m above sea level, and also at the sites of the Nicaraguan tsunamis, where most of the trees survived the tsunami flows, despite being swamped. All the signs indicate that the tsunami flow forces at Riangkrok must have been fiercely strong, as they ripped away everything that had stood in the area. At this small rural village, 137 people lost their lives to the tsunami.

Harry Yeh

Harry Yeh is in the Department of Civil Engineering, University of Washington, Seattle, Washington 98195, USA.

frequencies (5 mHz) and at those typical of tsunami waves (1 mHz). Kanamori and Kikuchi have exceptionally clear evidence that the Nicaraguan earthquake arose from a very slow rupture. Nevertheless, its true size (or moment) was accurately given within the hour following the event, as was the case for the Indonesian shock three months later.

For some earthquakes, no clue that large tsunamis were imminent can be sifted from the seismic records. For such events as the 1946 Aleutian tsunami (which destroyed the city of Hilo, Hawaii), it has been proposed that there was massive failure of the sea floor in the form of underwater slumping of sedimentary masses during or after the rupture⁷. Precedents can be found on land (such as the explosion of Mount St Helens in 1980), or even at sea, in the case of the large earthquake off the Grand Banks of Newfoundland in 1929⁸.

Kanamori and Kikuchi reconcile the two kinds of tsunami earthquakes, at least in part, by suggesting that sedimentary structures are crucial in both situations. On the one hand, they can account for the slowing down of the seismic rupture, when the latter is forced into mechanically inferior materials such as subducted sedimentary layers; rupture in sedimentary layers does enhance tsunami generation⁹. This mechanism could explain the large Nicaraguan tsunami, which took place in an area where the sparse sedimentary cover of the young plate is subducted straight into the trench without being scraped and piled up as an accretionary structure. On the other hand, where large sedimentary structures exist on the ocean floor, a sizeable earthquake can lead to their total failure in an episode of slumping, providing exceptional energetic coupling with the water column, and throwing up a large tsunami wave.

Local slumps are also thought to be important in smaller earthquakes, such as the 1989 Loma Prieta shock in California. Despite its epicentre being on land, this shock generated a small tsunami in Monterey Bay through a small landslide triggered by strong underwater motion¹⁰. It is too early to say for sure what caused the exceptional tsunami am-

plitudes on 12 December at Riengkrok, but the earthquake certainly resulted in subaerial landslides. One may shiver at the thought of what was perhaps the grand-daddy of all Pacific tsunamis, when prodigious underwater slumps, each involving up to 5,000 km³ of sediment, allegedly took place a few million years ago along the Hawaiian chain¹¹.

The goal of research in tsunami warning can be stated simply as the eradication of tsunami earthquakes, if we follow Kanamori's definition of these as events whose tsunamis we do not understand on

the basis of their seismic waves. Kanamori and Kikuchi have made a considerable step forward by digging deeper into the nature of the seismic source and the origin of its low-frequency components. We may realistically hope that slow earthquake components will in the future be identified in time to give coastal residents the signal to run for the hills. □

Emile A. Okal is in the Department of Geological Sciences, Northwestern University, Locy Hall, 1847 Sheridan Road, Evanston, Illinois 60208, USA.

VIRAL TRANSACTIVATION

Pleiotropy and henchman X

Michael Gilman

STANDARD dime-novel detective practice is to tail the henchman; eventually, he will lead you to the master criminal. So it is with virologists, who have reasoned that studying regulatory proteins, especially those of small viruses heavily dependent on host-cell functions, will lead them to the master regulatory machinery of the cell. But these proteins have not yielded their secrets easily, in large part because their activity can result in a variety of consequences. This is pleiotropy, and disentangling the many functions of pleiotropic activators has been the principal challenge facing virologists.

The report by Kekulé *et al.* on page 742 of this issue¹ may lead to the unravelling of pleiotropy for a particularly perplexing viral transactivator, the X protein of hepatitis B virus. It seems that X (denoted HBx in the paper) activates protein kinase C, a key component of cellular signal transduction; in doing so, it achieves a pleiotropic effect on many transcription factors that respond to extracellular signals, and presumably prepares the host cell for the task of replicating the viral genome. Moreover, by targeting protein kinase C, X mimics the activity of tumour promoters, which may explain the virus's unusual oncogenic properties.

Hepatitis B virus (HBV) is a small DNA virus with a genome of 3.2 kilobases which sports a mere four open reading frames. One of these encodes the regulatory protein X, with a predicted relative molecular mass of 17,000. The gene for X is conserved in hepadnaviruses and, where it can be examined, seems to be required for viral replication. The X protein is a transcriptional activator of both the HBV genome and a panoply of cellular promoters²⁻⁴. In particular, it activates reporter genes under the control of signal-responsive elements such as AP-1/CRE sites and kB elements⁵⁻⁷.

Kekulé *et al.* now show that activation of AP-1 by X is blocked by several specific inhibitors of protein kinase C (PKC) and by downregulation of the enzyme. They also find that transient expression of the X gene causes production of the membrane-associated lipid activator of PKC, diacylglycerol, and rapid translocation of PKC from cytosol to membrane, a hallmark of PKC activation. Thus, under these conditions, X appears to act (directly or indirectly) by triggering a phospholipase, mimicking an activity of growth factors and hormones.

Although this observation can neatly explain how X can activate so many promoters, the literature remains replete with conflicting results. Two groups have reported that X can function as a transcriptional activator when targeted to a promoter by fusion to a heterologous DNA-binding domain^{5,8}. And the protein appears to interact directly with members of the CREB/ATF transcription factor family⁹. Together these data suggest that X has a rather more orthodox transactivation function, one that would not seem to involve changes in phospholipid metabolism. Clearly, however, nearly all the transcription factors affected by X are also responsive to cellular signal transduction pathways, so the involvement of the protein in signal transduction may be the common thread that links many of these observations.

What, then, is the direct target of protein X action? It could be a phospholipase, but this is not sufficient to account for all of the protein's reported activities. More likely, X is itself pleiotropic, either because it directly modifies several target proteins, perhaps through an associated protein kinase activity¹⁰, or because, like a growth-factor receptor, it activates a number of cellular signal-transduction pathways. This latter possibility is attractive because it could account for the observation that X can

1. Kanamori, H. & Kikuchi, M. *Nature* **361**, 714–716 (1993).
2. Kanamori, H. & Given, J. W. *Phys. Earth planet. Inter.* **27**, 8–31 (1981).
3. Talandier, J. & Okal, E. A. *Bull. seism. Soc. Am.* **79**, 1177–1193 (1989).
4. Satake, K. *Proc. 2nd US–Japan Tsunami Workshop*, Univ. Hawaii, Honolulu, 6–9 November 1990.
5. Talandier, J. & Okal, E. A. *Pure appl. Geophys.* **138**, 43–60 (1992).
6. Kanamori, H. *Phys. Earth planet. Inter.* **6**, 346 (1972).
7. Kanamori, H. *Proc. XXIIIrd Gen. Assembly IASPEI*, Tokyo, 19–30 August, 1985.
8. Hasegawa, H. S. & Kanamori, H. *Bull. seism. Soc. Am.* **77**, 1984–2004 (1987).
9. Okal, E. A. *Natural Hazards* **1**, 67–96 (1988).
10. Ma, K.-F., Satake, K. & Kanamori, H. *Geophys. Res. Lett.* **18**, 637–640 (1991).
11. Moore, J. G. *et al. J. geophys. Res.* **94**, 17465 (1989).

activate kB elements without a requirement for PKC (ref. 7).

Perhaps the most interesting aspect of the results reported by Kekulé *et al.* is how they bear on oncogenesis by HBV. This virus is a serious medical problem, infecting some 200 million people; it is strongly associated with the development of liver cancer, but the latency of the disease is long and there has been no consistent molecular *modus operandi* for how the virus contributes to tumour formation. In this regard, HBV bears a striking resemblance to HTLV-1, which also harbours an inscrutable transactivator, *tax* (which once also bore the name X). Both viruses cause tumours with long latencies, and although neither X nor *tax* acts as a conventional oncogene, these regulatory genes and their subtle effects on cellular gene expression almost certainly make a material contribution to transformation. The X protein has been directly tested for its oncogenic potential in transgenic mice with mixed results, and it does not transform NIH3T3 cells¹¹. But by activating PKC, X may mimic compounds such as phorbol esters, which are powerful tumour promoters — that is, rather than delivering the blow that converts a hepatocyte into a cancer cell, X may perturb cell growth more subtly, expanding the target for subsequent genetic events or favouring the outgrowth of populations that have already sustained genetic damage.

But the case of henchman X is not closed. Too many inconsistencies persist, and too much pleiotropy remains unexplained. Still, the results reported by Kekulé *et al.* shift the focus of our thinking about X. The cellular machinery implicated by X is, for the moment, not the transcriptional machinery nor even a cyclin-dependent protein kinase. It is instead the signal transduction apparatus of the cell, for by perturbing the flow of information in the cell, X can affect the expression of many genes and the growth state of the cell itself. □

Michael Gilman is at Cold Spring Harbor Laboratory, PO Box 100, Cold Spring Harbor, New York 11724, USA.

Seeing the mantle in the round

Scott D. King

CAN material be transported between the upper and lower mantle? This question has exercised geochemists and geophysicists for years; without knowing the answer, they cannot infer the Earth's thermal and chemical history. But the computational task is immense. The properties of the mantle must be pared down to their essentials, and modellers then have the frustrating task of deciding which of these properties — compressibility, rheology, spherical geometry, changes in mineral phase — must be sacrificed. On page 699 of this issue¹, Tackley and his co-workers show that convection calculations in three-dimensional spherical geometry combined with an endothermic phase transformation at 670 km depth may resolve long-standing questions about large-scale structure and flow patterns in the Earth's interior.

The centre of attention is the 670-km seismic discontinuity, a region of the mantle where seismic velocities increase by about 6 per cent over a few kilometres of depth. Laboratory experiments have shown that the endothermic phase transformation from one of the main mantle constituents, the γ -spinel phase of olivine ($(\text{Mg,Fe})_2\text{SiO}_4$), to perovskite ($(\text{Mg,Fe})\text{SiO}_3$) plus magnesio-wüstite ($(\text{Mg,Fe})\text{O}$) occurs at a pressure that is compatible with the seismic observations.

Such a phase transformation can have important dynamical influences on mantle convection, and it will occur at greater depths in a cold, downgoing slab than in the average ambient mantle. The depression of the phase transformation creates a density anomaly in the slab which opposes the thermal density anomaly, reducing the net downward gravitational force on the slab. It is not hard to see that, given a large enough density anomaly from the phase transformation, the slab could become neutrally buoyant. The reverse is true for hot, upwelling plumes: the phase transformation will occur at shallower depths, providing a force that opposes the less-dense, rising plume material.

Conventional wisdom held that a change in chemical composition was required to enforce layered convection, because linear stability analysis indicated that an endothermic phase transformation would need an unrealistically large negative Clapeyron slope if it were to stop subducting material from sinking into the lower mantle². In fact, it now seems that the endothermic spinel-perovskite phase transformation can stratify a convecting fluid^{3–5}.

In studying the mantle, 3D calculations are an essential rather than a luxury. Without them, it is tricky to relate the complex, time-dependent behaviours observed in 2D calculations to experimental observations of a 3D mantle. In the 2D axisymmetric spherical calculations^{6,7}, a cyclic behaviour was seen: cold downwelling material accumulated above the phase transformation up to a critical amount and would then drain into the lower mantle, changing the large-scale flow pattern. This led to speculation^{6,7} that the break-up of the ancient supercontinents and reorganization of tectonic plates worldwide were linked to the dynamics of the 670-km phase transformation. Tackley and co-workers observed similar periodic events, which they dub “flushing events”, in their 3D calculations. But these events are more localized, and at any given time, several such flushes are likely to be in progress in the mantle, suggesting that they are less catastrophic than the events predicted from the 2D calculations. Presumably it is the extra degree of freedom in the 3D calculations that allows the flushing events to occur with less dramatic consequences for the global circulation pattern.

The (much disputed) final fate of subducting slabs is the key to the question of mantle layering. Residual sphere and regional tomography studies have suggested that slabs may penetrate the 670-km discontinuity in the northwest Pacific⁸, whereas they are bent and slide horizontally along the discontinuity under the western Pacific⁹. Tackley and his co-workers find that the cold downgoing sheets in their calculations generally do not penetrate the phase transformation, as they do not have the necessary density contrast. When the

1. Kekulé, A. S., Lauer, U., Weiss, L., Luber, B. & Hofschneider, P. H. *Nature* **361**, 742–745 (1993).
2. Tzu, J.-S. & Schloemer, R. H. *J. Virol.* **61**, 3448–3453 (1987).
3. Spandau, D. F. & Lee, C.-H. *J. Virol.* **62**, 427–434 (1988).
4. Zahn, P., Hofschneider, P. H. & Koshy, R. *Oncogene* **3**, 169–177 (1988).
5. Seto, E., Mitchell, P. J. & Yen, T. S. B. *Nature* **344**, 72–74 (1990).
6. Mahé, Y. *et al.* *J. biol. Chem.* **266**, 13579–13763 (1991).
7. Lucito, R. & Schneider, R. J. *J. Virol.* **66**, 983–991 (1992).
8. Unger, T. & Shaul, Y. *EMBO J.* **9**, 1889–1995 (1990).
9. Maguire, H. F., Hoeffler, J. P. & Siddiqui, A. *Science* **252**, 842–844 (1991).
10. Wu, J. Y. *et al.* *Cell* **63**, 687–695 (1990).
11. Kim, C.-M., Koike, K., Saito, I., Miyamura, T. & Jay, G. *Nature* **351**, 317–320 (1991).

1. Tackley, P. J., Stevenson, D. J., Glatzmaier, G. A. & Schubert, G. *Nature* **361**, 699–704 (1993).
2. Schubert, G., Yuen, D. A. & Turcotte, D. L. *Geophys. J. R. astr. Soc.* **42**, 705–735 (1975).
3. Machetel, P. & Weber, P. *Nature* **350**, 55–57 (1991).
4. Peltier, W. R. & Solheim, L. P. *Geophys. Res. Lett.* **19**, 321–324 (1992).
5. Zhao, W., Yuen, D. A. & Honda, S. *Phys. Earth planet. Inter.* **72**, 185–210 (1992).
6. Peltier, W. R. & Solheim, L. P. *Eos (abstr.)* **73**, 43517 (1992).
7. Machetel, P. & Burnet, D. *Eos (abstr.)* **73**, 43518 (1992).
8. Creager, K. C. & Jordan, T. C. *J. geophys. Res.* **91**, 3573–3589 (1986).
9. Zhou, H.-W. *Phys. Earth planet. Inter.* **61**, 199–229 (1990).
10. Tanimoto, T. *Geophys. J. int.* **100**, 327–336 (1990).
11. Su, W. & Dziewonski, A. M. *Nature* **352**, 149–152 (1991).
12. Jarvis, G. T. & Peltier, W. R. in *Mantle Convection: Plate Tectonics and Global Dynamics* (ed. Peltier, W. R.) 479–594 (Gordon and Breach, New York, 1989).
13. Hansen, U., Yuen, D. A. & Kroening, S. E. *Phys. Fluids A2*, 2157–2163 (1990).
14. Ita, J. & Stixrude, L. J. *geophys. Res.* **97**, 6849–6866 (1992).

volume of subducting material reaches a critical amount, however, it initiates a flushing event, which drains the subducted material into the lower mantle. In this way, the seismic observations can be reconciled.

The endothermic phase change could explain another paradox in the mantle. The lateral seismic-velocity structure seen in tomographic images of the mantle^{10,11} is dominated by long wavelengths; this is not what would be expected from calculations of high-Rayleigh-number convection in a uniform fluid^{12,13}. But when we add the endothermic phase transformation to 3D calculations, the downgoing flow accumulates in the upper mantle, creating large, spatially coherent anomalies. This new feature could account for the long-wavelength thermal anomalies.

Although the lack of continents and plates on the surface of the modelled spherical shell means that the mantle still retains some mysteries, Tackley and his colleagues have demonstrated that 3D calculations of mantle convection in the Rayleigh number range appropriate for the Earth are now truly possible. Numerical investigations in 3D spherical geometry are important to tie in dynamical predictions with observations, although it would be good to investigate thermodynamic properties closer to those of the real mantle¹⁴. The greater challenge is to add laterally varying viscosity to 3D calculations. □

Scott D. King is in the Department of Earth and Atmospheric Sciences, Purdue University, West Lafayette, Indiana 47907, USA.

CONSERVATION GENETICS

Breeding like flies

Katherine Ralls and Robin Meadows

BREEDING rare and endangered animals in captivity is like gambling — the stakes are high and the outcome is uncertain. Although theories for breeding animals in captivity are widely accepted and applied, there has been little experimental support for them because of the small numbers of animals involved, their long generation times and the expense of maintaining them. Now a team of Australian researchers led by Richard Frankham provides both welcome confirmation of current captive breeding strategies and some disturbing surprises¹⁻⁵.

Scientific consensus holds that captive endangered species such as the golden lion tamarin (see picture, over) must be managed to preserve genetic variation, which will maximize their reproduction and survival in captivity as well as their ability to adapt to future environmental changes when returned to the wild. Current breeding strategies designed to preserve genetic variation in small captive populations are rooted in population genetics theory^{6,7}. Because this theory is based on a simple model that fails to account for natural selection, genetic mutation and linkage, and other phenomena⁸, it is important to test the theory's predictions experimentally.

To circumvent the problems of using vertebrates, the Australian team used fruit flies (*Drosophila melanogaster*), which can be raised in great numbers easily and inexpensively. Moreover, *Drosophila* have been well characterized genetically and they are good genetic models for vertebrates⁹⁻¹¹.

The experiments began with an out-

bred *Drosophila* population founded with 202 inseminated females. Although the exact design varied from experiment to experiment, in general several replicate populations were managed according to a specific breeding strategy and their genetic diversity and fitness then compared with those of a randomly bred control population. Measures included genetic variation, which was gauged by allozyme electrophoresis; mean inbreeding coefficient, which was calculated from population pedigrees; and reproductive fitness after several generations of management, which was defined as the number of progeny from the managed *Drosophila* relative to the number from a chromosomally marked baseline strain.

As predicted by population genetics theory, the experiments show that loss of genetic variation in captive populations is slowed by three strategies: equalizing the genetic representation of each wild-caught animal used to found the descendant captive population¹, equalizing the sizes of the families derived from each animal² and equalizing the sex ratio of breeding adults (R. Frankham, personal communication).

Equalizing the representation of founders is necessary because some usually produce more offspring than others, making their genetic contributions to the descendant population unequal. Founder representation can be equalized in subsequent generations by limiting the reproduction of individuals descended from prolific founders and encouraging the reproduction of those descended from less prolific founders¹.

RÉSUMÉ

Fishy business

THE discovery of a new kind of antibiotic in the skin of frogs followed when it struck M. Zasloff that *Xenopus* never caught infections when tossed back into the tank after surgery. This set in train a search among the obscurest animal tissues, which turned up alkaloids, lipids and peptides with interesting bactericidal properties. The latest such find reported by Zasloff and colleagues (K. S. Moore *et al.*, *Proc. natn. Acad. Sci. U.S.A.* **90**, 1354–1358; 1993) occurs in the stomach of a shark. Termed squalamine, it is a sulphated bile salt, conjugated to spermidine, and is highly active against Gram-positive and Gram-negative bacteria; it is, moreover, fungicidal and induces lysis of protozoa. The authors speculate that squalamine acts as a systemic antimicrobial agent in the shark, but how it might do so is a mystery.

Winter warmer

PARTS of North America and northern Eurasia may have been warmer in winter as an indirect consequence of volcanic eruptions (A. Robock and J. Mao, *Geophys. Res. Lett.* **12**, 2405–2408; 1992). Looking through temperature records back to 1883, they found that all 12 of the largest eruptions were followed by unusually mild winters in those parts, but colder than average winters in the Middle East. The pattern developed in the first winter following eruptions in the tropics, but could be delayed until the next year for eruptions at higher latitudes. It seems that absorption of sunlight by volcanic aerosols warms the tropical stratosphere, strengthening the Equator-to-Pole temperature gradient; the stronger winds that ensue bring warmer marine air over the continents. On average, though, the overall effect of an aerosol shadow is to lower ground temperatures globally for several years.

Risk perception

THE consequences of an epidemiological study that went off at half-cock are described by T. L. Guidotti and P. Jacobs in the *American Journal of Public Health* (**83**, 233–239; 1993). Early in 1987, a draft report showing a 25% excess of cancers in two suburbs of Edmonton, Alberta, received wide currency. The report was in error, but before that was made public Guidotti and Jacobs enlisted the help of an investigative reporter and carried out surveys of a variety of responses to the perceived hazard. People seem in general to have reacted in a remarkably robust manner, though house prices fell by some 5%. The cause of the error, it turned out, was that the population of one suburb had not been counted in the 'population at risk', though its cancer incidence had been.

In practice, however, strict equalization is sometimes inappropriate. For example, if a founder has only one offspring and then dies, only half of its genes survive.

Similarly, family sizes can be equalized by allowing each animal in a captive population to contribute the same number of offspring to the next generation. The Australian team found that equalizing family sizes effectively halved the

introducing new animals to captive populations³ — introduction of a single immigrant from one partially inbred population to another increased the second population's reproductive fitness from 0.29 to 0.63 (relative to a fitness of unity in the outbred base population). Unfortunately, this strategy is not an option for captive populations of California condors (*Gymnogyps californianus*), black-footed ferrets (*Mustela nigripes*) and other species that are extinct in the wild.

Jessie Cohen/National Zoological Park

Another sobering finding is that captive animals rapidly adapt genetically to captivity⁴. Animals adapted to captive conditions are likely to reproduce more poorly in the wild, meaning that captive endangered species should be returned to the wild as soon as possible. Furthermore, the observations suggest that even very large captive populations may lose genetic variation at rates greater than expected⁵. For example, a population of 1,000 *Drosophila* lost 86 per cent of its genetic variation in 2.5 years, making its effective population size much smaller than its census size. This finding is supported by independent work on fish¹².

Overall, the implications of these studies with *Drosophila* are twofold — genetic management is necessary even for

very large captive populations, and many small populations may be even more endangered than is generally recognized. Perhaps even more importantly in the long run, they show that work in the laboratory has a great deal to offer in the validation and refinement of conservation genetics theories. □

Katherine Ralls is at the National Zoological Park, Smithsonian Institution, Washington, DC 20008, USA. Robin Meadows is at 460 Crofters Court, Fairfield, California 94533, USA.

1. Loebel, D. A., Nurthen, R. K., Frankham, R., Briscoe, D. A. & Craven, D. *Zoo Biol.* **11**, 319–332 (1992).
2. Borlase, S. C. et al. *Conserv. Biol.* (in the press).
3. Spielman, D. & Frankham, R. *Zoo Biol.* **11**, 343–351 (1992).
4. Frankham, R. & Loebel, D. A. *Zoo Biol.* **11**, 333–342 (1992).
5. Briscoe, D. A. et al. *Conserv. Biol.* **6**, 416–425 (1992).
6. Frankel, O. H. & Soule, M. E. *Conservation and Evolution* (Cambridge University Press, 1981).
7. Ralls, K. & Ballou, J. (eds) *Zoo Biol.* **5**, 81–238 (1986).
8. Wright, S. *Evolution and Genetics of Populations* Vol. 2, *The Theory of Gene Frequencies* (University of Chicago Press, 1969).
9. Robertson, A. *Proc. X Int. Congr. Genet.* **1**, 199–205 (1959).
10. Hammond, K. *Proc. 1st World Congr. Genet. Appl. Livestock Prod.* **1**, 425–437 (1974).
11. Frankham, R. *Proc. 2nd World Congr. Genet. Appl. Livestock Prod.* **5**, 43–56 (1982).
12. Bartley, D., Bagley, M., Gali, G. & Bentley, B. *Conserv. Biol.* **6**, 365–375 (1992).

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Golden lion tamarins (*Leontopithecus rosalia*) — as with many other species, captive breeding programmes may be essential for their survival.

captive breeding space required to maintain a given degree of genetic diversity². This implication matters because there is a great shortage of captive breeding space for endangered species⁷.

Another way to increase the effective size of the captive breeding space is to equalize the breeding sex ratio (R. Frankham, personal communication). Allowing equal numbers of males and females to contribute to the next generation is easy in species that are monogamous, but difficult in those that are polygynous; most mammals and some birds are polygynous. Although artificial insemination is a potential solution, the appropriate techniques have not yet been developed for most endangered species. Available approaches include keeping fewer females with each breeding male and replacing breeding males frequently.

What of the disturbing surprises? The experiments show that even 'properly managed' populations of captive *Drosophila* lost 74 per cent of their reproductive fitness after 11 generations and had lower genetic diversity than large wild populations. Frankham and his colleagues found that this decrease in reproductive fitness can be mitigated by

DAEDALUS

The nuclear mole

THE melt-down of a nuclear reactor has been dubbed 'The China Syndrome', after the country towards which a melting-down Western reactor would head, as it sank white-hot into the Earth. What an excellent method, says Daedalus, of getting rid of a reactor! How much cheaper than painfully decommissioning and dismantling the radioactive hulk! Sadly, modern reactors are not designed as geological submarines. Attempts to melt them down would throw a lot of radioactive material into the atmosphere.

This method of disposal would be better applied to high-level radioactive waste, which generates heat more predictably. Daedalus envisages a dense mass of such waste confined in a container of some high-melting, inert material such as tungsten or tantalum carbide, and left to melt its way steadily down into the Earth.

The waste need not even be ferociously hot. Daedalus reckons that it should be lowered down a very deep borehole, geothermally heated at the bottom already, which would then be filled with water. Pressurized by the hydrostatic head and heated further by the radioactive capsule, the water at the bottom would become supercritical. Supercritical water is an excellent solvent for silica and silicate rocks, so the capsule would dissolve its way into the depths. The shaft above it would slump shut and be sealed by recrystallized rock, trapping the water and raising its pressure further. In due course the capsule would reach rock so hot that it could progress by actual melting. Soon it would be sinking freely through already molten rock. Ultimately, its refractory container would itself melt. The deadly contents would dissolve and disperse safely, hundreds of kilometres down in the Earth's mantle.

Daedalus's 'nuclear mole' could also act as a useful deep-Earth probe. With luck, the resolidified channel left behind it would act as an acoustic fibre or sound-pipe, transmitting seismic pulses without the attenuation of a freely spreading signal. Impulses launched from the surface would then be reflected back by each discontinuity in the channel, as well as by the probe itself. Their travel time and doppler shift would establish the depth and sinking speed of the probe, and the acoustic character of the strata through which it had passed. If the probe or some sensor on it had a resonant frequency sensitive to temperature and pressure, it could relay these conditions too. Even when the probe itself had melted, the acoustic channel would remain, a sensitive geological stethoscope transmitting intimate rumbles from the very bowels of the Earth.

David Jones

Dental HIV transmission?

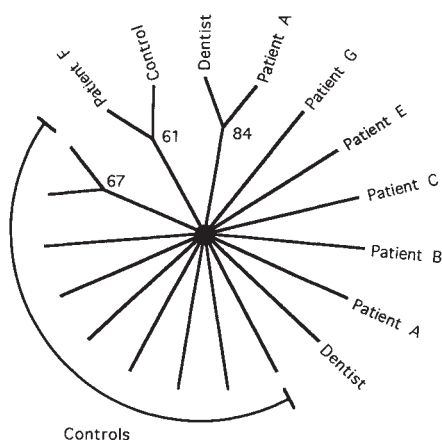
SIR — It is frequently reported, on the basis of the conclusion reached by Ou *et al.*¹, that a Florida dentist infected five of his eight HIV-1 seropositive patients. The molecular data, consisting of nucleotide sequences from the highly variable C2–V3 region of the HIV-1 envelope gene, are particularly crucial because the epidemiological investigation was retrospective: all the patients were aware of the dental transmission hypothesis². Because this case remains unique in terms of the hypothesized

We tested these hypotheses using new sequences from the dental patients and a new set of regional controls. Ou *et al.* selected four controls that were similar to the dentist's sequences for detailed analysis, and we selected ten controls. This selection is justified: the dental group should be monophyletic compared to any controls. In fact, the test is biased in favour of accepting the dental transmission hypothesis because the controls in both studies were obtained at clinics about 90 miles from the dentist's practice area.

Although parsimony analysis gives a monophyletic dental group, some trees that mix controls and the dental group require only one additional substitution. One weakness of parsimony analysis is the assumption that all nucleotide changes are approximately equally likely. This is unlikely to be the case for the C2–V3 region; many sites are identical across North America, whereas other sites commonly vary within individuals. A slightly different analysis, threshold parsimony, allows for such a situation³. All the best threshold parsimony trees include one or more controls within the dental group, whether our data or those of Ou *et al.* are used.

Parsimony analysis yields a point estimate, whereas the hypothesis test requires an estimate of confidence intervals. Ou *et al.* used bootstrap resampling and reported finding a monophyletic dental group in 79% of the replicates. In contrast, we find a monophyletic dental group in only 59% of the replicates, and in fewer than 50% of the replicates using threshold parsimony, a result due to the greater similarity of our controls to the dental group (see figure). The exact level required for significance in the bootstrap test is not known in general, but in the simplest case (four sequences and a molecular clock) greater than 80% is required, and values less than 75% mean that the relationship in question cannot be resolved⁴.

Ou *et al.* used two additional methods, but neither provides a direct test of the dental transmission hypothesis. The Wilcoxon rank-sum test is inappropriate because it compares average properties, not individual properties, of the patient and control groups. The other test used a pattern of eight noncontiguous amino acids said to form the dentist's 'signature'. In fact, none of the signature residues is unique to the dental group; one of our controls has six of the eight, and five have four or more. Further, examining the signature as phylogenetic data reveals that seven of the sites provide evidence for groups containing both patients and controls.



Statistical analysis of C2–V3 sequence data: 100 bootstrap replicates were analysed by threshold parsimony (threshold value 2.0). Very few nodes were found in more than 50% of the replicates, indicating that there is little reliable phylogenetic information in these data (no inference about relationships should be drawn from the relative positions of branches coming from the central node). The 'dental group' consists of patients A, B, C, E and G. Additional details can be obtained from R. W. D.

mode of virus transmission, the conclusion and analyses used to support it should be carefully examined.

What is the null hypothesis against which the dental transmission hypothesis should be tested? Population genetics suggests that a rapidly evolving marker can develop strong geographical substructure; therefore, an appropriate null hypothesis is that the patients independently acquired similar variants within the local community. These two hypotheses can be distinguished by the inferred historical relationships among the viral sequences. The dental transmission hypothesis requires that a branch on the viral phylogenetic tree lead to the dental group alone and not include any controls. In phylogenetic terms, the dental group must be monophyletic. The null hypothesis would be rejected if a tree with a monophyletic dental group is significantly better supported than any tree with controls intermixed within the dental group.

Our analyses show that the available data are consistent with both the dental transmission hypothesis and the null hypothesis and do not yet distinguish between the two. Much of the difficulty lies in the decision of Ou *et al.* to focus on the small, highly variable C2–V3 region. A more conclusive test of the hypothesis that these patients were infected with HIV during dental treatment will require a larger dataset from some other region of the HIV genome.

Ronald W. DeBry

Lawrence G. Abele

Department of Biological Science,
Florida State University,
Tallahassee,
Florida 32306, USA

Stanley H. Weiss

Division of Infectious Diseases
Epidemiology,

Department of Preventive Medicine and
Community Health,

New Jersey Medical School,
Newark,

New Jersey 07107, USA

Martin D. Hill

Marla Bouzas

Eric Lorenzo

Folke Graebnitz

Lionel Resnick

Retrovirology Research Laboratories,
Mount Sinai Medical Center,
Miami Beach,
Florida 33140, USA

1. Ou, C.-Y. *et al.* *Science* **256**, 1165–1171 (1992).
2. Ciesielski, C. *et al.* *Ann. intern. Med.* **116**, 798–805 (1992).
3. Felsenstein, J. *Biol. J. Linnean Soc.* **16**, 183–196 (1981).
4. Zharkikh, A. & Li, W.-H. *Molec. Biol. Evol.* **9**, 1119–1147 (1992).

Rose is a rose...

SIR — I apologize to your readers for any confusion caused by my deducing the names of genes from those of mutants¹, thank Enoch *et al.* for correcting me², and blame Gertrude Stein for suggesting the idea that a mutant names a gene names a protein³. Finally, in an effort to minimize confusion about the *hus* mutants, I note that their names appear to have changed between Enoch *et al.*'s first⁴ and second⁵ publications on the subject (for example *hus14* in ref. 4 became *hus1-14* in ref. 5).

Andrew Murray

Department of Physiology,
University of California,
San Francisco,
California 94143-0444, USA

1. Murray, A. W. *Nature* **359**, 599–604 (1992).
2. Enoch, T., Carr, A. & Nurse, P. *Nature* **361**, 26 (1993).
3. Stein, G. *Sacred Emily* (Something Else, New York, 1968).
4. Enoch, T., Gould, K. L. & Nurse, P. *Cold Spring Harb. Symp. quant. Biol.* **56**, 409–416 (1991).
5. Enoch, T., Carr, A. M. & Nurse, P. *Genes Dev.* **6**, 2035–2046 (1992).

Hox genes and serial homology

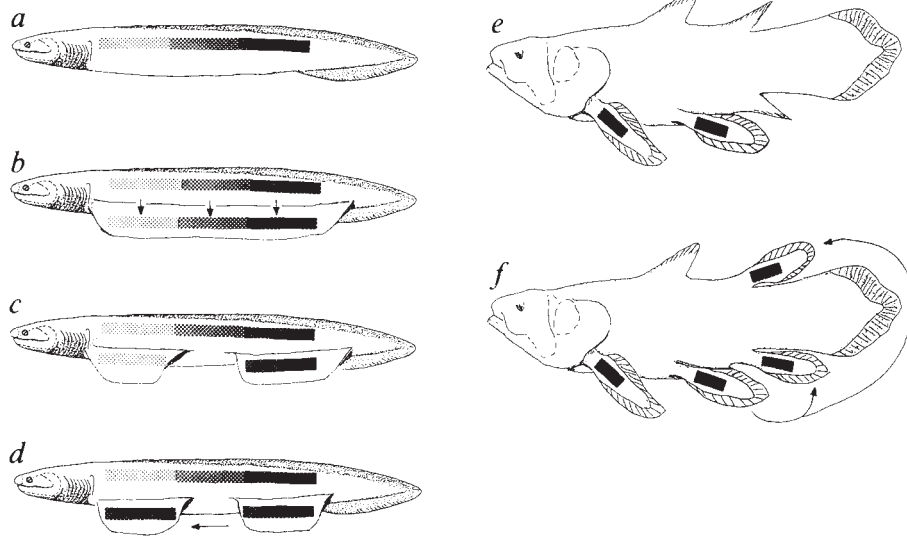
SIR — Fore and hind limbs of land-dwelling vertebrates have very similar structures despite having been independently derived from different fins of fish. We suggest that *Hox* genes, which are now expressed both in developing fore and in hind limbs, were originally expressed only in the hind limb and controlled its pattern. Their subsequent

lateral plate and somitic mesoderm adjacent to the fore limb bud than are expressed next to the hind limb bud². The genes of the *Hox C* cluster expressed in each embryonic appendage are different, and in each case are those expressed in the adjacent mesoderm^{3,4}. However, the genes of the *Hox A* and *Hox D* clusters expressed within both the developing limbs are the same, and are those that are expressed in the posterior lateral plate and somitic mesoderm of the main body axis. In addition, the *Hox A* and *Hox D* gene expression

Hox genes. In each case the *Hox* genes were used, as in the flank, to provide positional cues during embryological patterning. Subsequently, in the lineage leading to the tetrapods, the posterior genetic programme was co-opted for use in specifying the pectoral fin pattern. This was achieved by activating the posterior *Hox A* and *D* genes in place of the anterior *Hox* genes, perhaps by ectopically initiating signalling centres that might originally have been specific to the hind limb, such as the zone of polarizing activity. Because the embryonic domains of *Hox* gene expression have been shown to influence directly the pattern of bones in the limb⁸, this would explain the serial homology in the pattern of bones of tetrapod fore and hind limbs.

It is uncertain phylogenetically when the homeotic transformation took place. We emphasize that all the available *Hox A* and *Hox D* cluster evidence comes from tetrapods. The earliest known 'tetrapod' fossils from the Devonian are suggestive of a creature with leg-like hind limbs and fish-like pectoral appendages⁹. This might indicate that the alteration in *Hox* expression happened in connection with the origin of tetrapod limbs, and that the derived character state is unique to Tetrapoda. Alternatively, the transformation could have occurred earlier, perhaps at the time of the emergence of certain lobe-finned groups, such as the coelacanth, which have very similar pectoral and pelvic fins.

In most fish the paired pectoral and pelvic fins are quite different in form from the anal and dorsal fins. However, in coelacanth the anal and posterior dorsal fins are strikingly similar in morphology to the paired fins, particularly resembling the pelvis at their base¹. In the living representative of the coelacanth, *Latimeria*, the correspondence includes the internal skeletons, musculature and innervation, and also extends to the use of the fins in swimming. This was interpreted by Ahlberg as evidence for a homeotic transformation of the median fins to a paired fin form¹. Although the genetic basis for this transformation is unknown, this is precisely the type of homeotic conversion between appendages that we propose occurred at an earlier time between the pelvic and pectoral appendages. It seems likely that in both cases activation of the posterior



Model for the evolution of paired appendages. There are 13 *Hox* gene paralogue groups in four clusters expressed sequentially along the anterior–posterior body axis of modern vertebrate embryos, and in a similar pattern along the anterior–posterior axis of disparate phyla. It is supposed therefore that these genes were expressed in such a manner before the evolution of appendages, indicated by the bar with graded shading (a). According to current thought, the paired appendages evolved from a lateral fin fold which extended the length of the ancestral animal. It is proposed that the expression of *Hox* genes in the embryonic fin fold mirrored that in the axial mesoderm (b). The contraction of the fin fold into separate pelvic and pectoral appendages would have resulted in different *Hox* genes being expressed in the two fin buds, indicated by the bar with light shading in the anterior fin and the bar with darker shading in the posterior fin (c). A subsequent homeotic transformation at some point in the lineage leading to tetrapods results in the modern pattern of posterior *Hox* genes from two of the clusters being expressed in the pectoral appendage as well as the pelvic appendage (d). This would provide a genetic explanation for the serial homology in bone pattern between the fore and hind limbs. In the coelacanth lineage, the recently described homeotic conversion of the posterior dorsal fin and anal fin to a structure resembling the paired fins¹ is hypothesized to have also involved the activation of the posterior *Hox* genes (e and f).

ectopic activation in the developing fore limb resulted in a homeotic transformation, imposing the pattern of the hind limb on the fore limb. Support for our suggestion is provided by the analysis by Ahlberg in Scientific Correspondence¹ of coelacanth fins.

From phylogenetic comparisons we know that the *Hox* genes are an ancient gene family whose differential expression along the embryonic anterior–posterior body axis was used to provide distinct positional values long before the evolution of the vertebrate appendages. In tetrapods, different genes of the *Hox A*, *C* and *D* clusters are expressed in the

patterns in the fore limb are spatially and temporally almost identical to those in the hind limb^{5,6}.

The current view of the evolution of the paired appendages is that they are derived from a lateral fin fold which ran the length of an ancestral fish⁷. We suggest that the mesoderm that migrated laterally to form that primitive fin fold maintained expression of the *Hox* genes characteristic of their axial level in the adjacent flank. As the posterior region of the fin fold evolved into a pelvic fin it continued to express posterior *Hox* genes, while at the anterior the nascent pectoral fin initially expressed anterior

1. Ahlberg, P. E. *Nature* **358**, 459 (1992).
2. Kessel, M. & Gruss, P. *Cell* **67**, 89–104 (1991).
3. Oliver, G. et al. *Cell* **55**, 1017–1024 (1988).
4. Gada, A. M. C. et al. *Development* **116**, 497–506 (1992).
5. Dolle, P. et al. *Nature* **342**, 767–772 (1989).
6. Yokouchi, Y., Sasaki, H. & Kuroiwa, A. *Nature* **353**, 443–445 (1991).
7. Zangerl, R. *Chondrichthyes I. Handbook of Paleichthyology* (ed. Schultze, H.-P.) 32 (Fischer, Stuttgart, 1981).
8. Morgan, B. A. et al. *Nature* **358**, 236–239 (1992).
9. Ahlberg, P. E. *Nature* **354**, 298–301 (1991).

Hox genes was a cause of the transformation. Theoretically, in *Latimeria* this could be tested. Future comparative studies on *Hox* gene expression in tetrapod, fish and shark embryos will shed light on when in the evolution of tetrapods the posterior *Hox* genes were first expressed in the pectoral appendage, and help to establish whether a shift in *Hox* expression is indeed the basis for the serial homology of fore and hind limbs.

Cliff Tablin

Ed Laufer

Department of Genetics,
Harvard Medical School,
200 Longwood Avenue,
Boston,
Massachusetts 02115, USA

Noble gases in atmospheres

SIR — Owen *et al.*¹ reported an interesting observation that noble gas isotope data for terrestrial samples (atmosphere, mid-ocean ridge basalt, Loihi volcanics), martian meteorites, and ice formed at 50 K lie on a line in a plot of $\log(^{84}\text{Kr}/^{132}\text{Xe})$ against $\log(^{36}\text{Ar}/^{132}\text{Xe})$ (see Fig. 2 of ref. 1). They suggested that the linear array represents a mixing line between two reservoirs with differing isotope abundances, one being the ice representing a comet and another an internal component in a planet. However, a line in a log-log plot does not represent mixing except for one very special case, where the line has a slope of unity. This specific condition clearly does not apply to the present case. Therefore, the mixing relation among Mars, Earth and comet noble gases

argued by Owen *et al.* must be dismissed.

The inadequacy of the mixing model can be further shown by comparing the noble gas isotope compositions between Earth and Mars. All the terrestrial materials (see figure) have identical noble gas isotope ratios (except for radiogenic ones)^{2,3}, whereas they are different from martian noble gas isotope ratios, especially in $^{38}\text{Ar}/^{36}\text{Ar}$. If the linear array for the terrestrial data is due to the mixing, the terrestrial data which are located close to Shergotty must have very similar isotope ratios to those of Shergotty, which is clearly not the case. Therefore, the above observation rules out the possibility that terrestrial noble gases are related to Mars noble gases by the mixing process proposed by Owen *et al.*¹.

We plotted newly obtained, high-quality data for mid-ocean ridge basalts and Loihi^{2,3}, and laboratory data for silicate melts and waters at various temperatures (0, 25, 50 °C) together with the Mars and ice data in a plot of $\log(^{36}\text{Ar}/^{132}\text{Xe})$ against $\log(^{84}\text{Kr}/^{132}\text{Xe})$ (see figure). The terrestrial data show a clearly linear array. Thus the linear array for the terrestrial samples must be attributed to cause(s) other than the mixing. It is known that equilibrium partition of noble gases between melt (either in silicate melt or in water) and gas gives rise to elemental (not isotope) fractionation to result in a linear relation in a plot of $\log(a/b)$ against $\log(c/b)$, where a , b and c stand for any of four (helium is non-conservative in the Earth) noble gases (see ref. 4). This is because noble gas solubility in the melt is, to a good approximation, proportional to $\exp(-kr_i^2)$ (where k is a numerical constant, r_i the radius of a noble gas atom

i), and hence the logarithm of elemental abundance ratio of noble gas a and b which is expressed as $\exp\{-k(r_b^2 - r_a^2)\}$ necessarily shows linear relation in a logarithmic plot (for detailed discussion, see ref. 5). Fractionation during gas-loss or gas-adsorption can also give rise to a similar linear trend to that seen in the figure (ref. 5). The rough linear relation among terrestrial samples in the figure can be attributed to one or some of these fractionation processes.

Although martian and terrestrial noble gases cannot be related to each other by the simple mixing process suggested by Owen *et al.*, the almost identical trend in the noble gas elemental abundance pattern both for Earth and Mars is remarkable, and deserves further investigation. Also, future investigation on noble gas isotope compositions on comets should resolve the significance of comets as a noble gas source for terrestrial planets.

Minoru Ozima

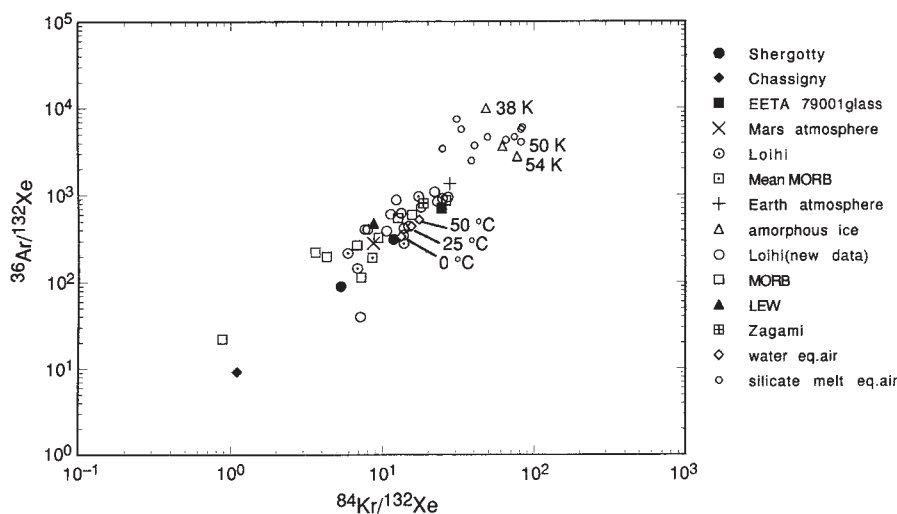
Naoko Wada

Department of Earth and Space Science,
Faculty of Sciences,
Osaka University,
Toyonaka City 560, Japan

OWEN AND BAR-NUN REPLY — Dr Ulrich Ott (Max Planck Inst. für Chemie, Mainz) has kindly pointed out to us that we erred in calling the straight line on our log-log plot a mixing line, representing the blending of two distinct volatile reservoirs. Ozima and Wada make the same observation above. The remarkable fit over three orders of magnitude in the value of $^{36}\text{Ar}/^{132}\text{Xe}$ and the similarity to a true mixing line over much of the distance had led us, the *Nature* referees, and various colleagues and audiences to overlook this rather obvious point.

We have therefore replotted all the data, making the simple assumption that there is a true mixing line that passes through the points for the atmospheres of Mars and Earth. In other words, we reassert our hypothesis that the heavy noble gases in the atmospheres of these two planets come primarily from the same two sources, one external and one internal to the planet. We have added data for the SNC meteorites Zagami⁶ and Lew 88516 (ref. 7). This plot is shown in the figure.

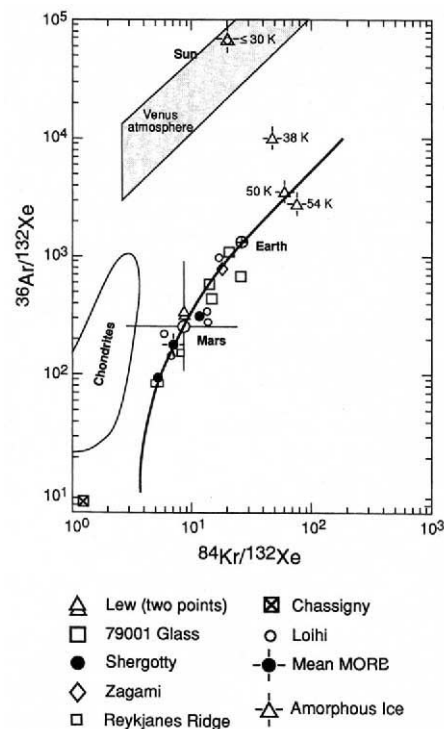
It is evident from this figure that the mixing line drawn under this assumption provides a reasonable fit to the data, again implicating ice formed at temperatures near 50 K as the major external contributor of these gases to the atmospheres of Mars and Earth. Only the SNC meteorite Chassigny is not accommodated by this hypothesis. We must therefore agree with Ott⁸ that the relationship (if any) between the gases in Chassigny



Terrestrial data (MORB (mid-ocean ridge basalt), Loihi volcanic rocks) and laboratory data (atmospheric noble gases absorbed in silicate melts such as basalt, andesite, ugandite, diopside, forsterite and in waters (0, 25, 50 °C)) are plotted in a $\log(^{36}\text{Ar}/^{132}\text{Xe})$ against $\log(^{84}\text{Kr}/^{132}\text{Xe})$ diagram together with Mars data. Scatter for the laboratory data (silicate melts) is probably due to the experimental uncertainty.

and the Martian atmosphere remains to be elucidated.

The discrepancy in $^{38}\text{Ar}/^{36}\text{Ar}$ to which we¹ and Ozima and Wada refer has only been found in EETA 79001 glass, not in Shergotty or the other SNCs (ref. 8). We believe that the observed linearity in the three isotope plot (see figure), which shows the Martian atmosphere coinciding with terrestrial rocks, argues for an



Data from our original Fig. 2b plus Zagami and Lew 88516 (see text) plotted with a true mixing line drawn through the points for the atmospheres of Mars and Earth.

atmospheric value of $^{38}\text{Ar}/^{36}\text{Ar}$ on Mars that is terrestrial (= meteoritic = cosmic).

We stress that we are pulling together various disparate data in the figure, representing several different sources and processes for the trapping of gases. For example, the SNC meteorites have had at least some fraction of the martian atmospheric gases they contain emplaced by shock⁹. Some of these gases could have been contributed by the impactor that dislodged the SNCs from Mars. (Note that some SNC data fall above the Mars atmosphere point in the figure, whereas all the terrestrial rocks appear below the Earth's atmosphere.) The trapping of gases in amorphous ice proceeds by adsorption in a microporous solid¹⁰. The gas-melt partitioning proposed for the terrestrial rocks by Ozima and Wada will not encompass these two cases, nor will it account for Chassigny because of the unique character of the latter's xenon isotopes, compared with other SNCs (ref. 8), nor will it explain the enhancement of radiogenic isotopes

in the martian atmosphere^{1,8}.

We find it highly significant that the data in the figure cluster around a mixing line that cannot be directly associated with either the chondritic meteorites or solar abundances (see refs 1, 8, 11). A mixture of gases contributed by comets formed in the Uranus-Neptune region of the solar nebula and gases from an internal, rocky reservoir offers an attractive alternative. The new noble gas data for submarine basalts^{2,3} support this suggestion. Of the 8 MORB and Loihi samples with $^{20}\text{Ne}/^{22}\text{Ne} \geq 10.2$, all have values of $^{36}\text{Ar}/^{132}\text{Xe} < 500$ (atmosphere values are 9.8 and 1,350, respectively). Partitioning from a common reservoir would not produce this association of isotopic (neon) and elemental (argon/xenon) fractionation, but the appearance of solar-type neon together with highly fractionated heavy noble gases in mantle-derived rocks (the internal reservoir) is exactly what our model predicts¹.

Nothing in this discussion explains the noble gases on Venus (see figure). If confirmed, however, a newly discovered¹² member of the long-sought Kuiper Comet-belt¹³ would be an example of the icy planetesimals that could account for the Venus abundances¹.

We continue to believe that our original hypothesis is viable, but it is clear that more data are required to test it. As we have said before^{1,14}, noble gas abundances and isotope ratios from dynamically new comets (as well as better data from Mars, Venus and submarine basalts) are essential.

Tobias Owen
Institute for Astronomy,
University of Hawaii,
Honolulu,
Hawaii 96822, USA

Akiva Bar-Nun
Department of Geophysics and
Planetary Sciences,
University of Tel-Aviv,
Tel-Aviv,
Israel

- Owen, T., Bar-Nun, A. & Kleinfeld, I. *Nature* **358**, 43–45 (1992).
- Hiyagon, H. et al. *Geochim. cosmochim. Acta* **56**, 1301–1306 (1992).
- Honda, M. et al. *Geochim. cosmochim. Acta* (in the press).
- Kirsten, T. *J. geophys. Res.* **73**, 2807–2810 (1968).
- Ozima, M. & Alexander, E. C. Jr *Rev. Geophys. Space Phys.* **14**, 385–390 (1976).
- Ott, U., Löhr, H. P. & Begemann, F. *Meteoritics (Abstr.)* **23**, 295–296 (1988).
- Ott, U. *Meteoritics (Abstr.)* **27**, 271 (1992).
- Ott, U. *Geochim. cosmochim. Acta* **52**, 1937–1948 (1988).
- Bogard, D. D., Hörz, F. & Johnson, P. H. *J. geophys. Res.* **91**, E99–E114 (1986).
- Bar-Nun, A., Kleinfeld, I. & Kochavi, E. *Phys. Rev.* **B38**, 7749–7754 (1988).
- Ott, U. & Begemann, F. *Nature* **317**, 509–512 (1985).
- Jewitt, D. & Luu, J. X. *Int. Astron. Union Circular* **5611** (1992).
- Kuiper, G. P. in *Astrophysics* (ed. Hynek, J. Allen) 357–424 (McGraw-Hill, New York, 1951).
- Owen, T., Bar-Nun, A. & Kleinfeld, I. in *Comets in the Post-Halley Era* (eds Newburn, R. Jr, Neugebauer, M. & Rahe, J.) 429–438 (Kluwer, Dordrecht, 1991).

More protein talk

SIR — Gonnet and Benner¹ have searched for the longest English word in the protein sequence databank. But what of other languages? Given that the ownership of the longest peptide-word will undoubtedly become a source of intense national pride, I thought it wise to investigate.

I performed a similar analysis to Gonnet and Benner using a standard hashing algorithm to search the SwissProt databank² with a multilingual word list of 1.3 million words from Danish, Dutch, English, Finnish, French, German, Italian, Norwegian, Spanish, Swedish and some Esperanto. Words were considered only if they incorporated no accented letters or other special characters.

Apart from English, four of the other languages provided nine-letter words: *ansvarlig* (a Danish word meaning 'liable') in entry HX_YEAST at position 85; *haletante* (French for 'breathless') in K1C0_XENLA at 145; *saltsilda* (Norwegian for 'salted herring') in PAI1_BOVIN at 271, and *stillassi* (the perfect subjunctive of the Italian word *stillare* — 'to drip') in STE2_YEAST at 207. The most apt nine-letter word was *salasivat*, PEHX_ERWCH at 10, the past tense of *salata*, Finnish for 'to keep hidden' or 'to encode'.

Although I did not find any other English words of nine letters or more in SwissProt, the search did turn up a 10-letter Italian word, *annidavate*, at position 45 in databank entry PHEA_FREDI (C-phycoerythrin α -chain from *Fremyella diplosiphon*). The word is from the past imperfect tense of the word *annidare*, meaning 'to nest'.

An honourable mention must also go to the 10-letter American-English word *Wallawalla* (the language of the Shahap-tian people of southeast Washington or Oregon), and the Dutch word *tariefklas* (literally 'tariff class') which are found reversed in the databank at position 18 of BVGB_BORPE and position 938 of ITA4_HUMAN, respectively.

The race is now on for the next longest word. How long will we have to wait before Germany finally scoops the honours with the possible 27-letter peptide-word for 'social sciences': *Gesellschaftswissenschaften*?

David Jones
Biomolecular Structure and Modelling
Unit,
Department of Biochemistry and
Molecular Biology,
University College,
London WC1E 6BT, UK

- Gonnet, G. H. & Benner, S. A. *Nature* **361**, 121 (1993).
- Bairoch, A. & Boeckmann, B. *Nucleic Acids Res.* **19**, 2247–2249 (1991).

Pernicious treatment

Fred S. Rosen

Lorenzo's Oil. A film directed by George Miller. Written by Nick Enright and George Miller. Released by Universal: 1993.

In 1912, when the Tsarevich of Russia had a nearly fatal haemorrhage at the imperial hunting lodge at Spala, his royal and distraught parents summoned Rasputin for consolation. The Holy One wired back from Siberia to the Tsarina and advised her to discharge the doctors. At the time, that was probably good advice, for there was nothing that medicine had to offer in the treatment of haemophilia. There was only the dimmest notion of the X-linked inheritance of the disease, although the Tsarina certainly knew that her brother and her maternal uncle, the Duke of Albany, had died from this bleeding disorder and that her sister, the Princess of Prussia, had two affected sons.

Some 70 years later, the Odone, less royal, equally distraught and more intelligent and hysterical than the Romanovs, discovered that their only male child had a fatal X-linked disease, adrenoleukodystrophy (ALD), for which modern medicine had no therapy. It was known that the disease is caused by an accumulation of very-long-chain fatty acids, resulting in progressive demyelination of the central nervous system and adrenal insufficiency. Michaela and Augusto Odone set out to find a dietary cure for their son, who was neurologically deteriorating quickly, and this is the subject of *Lorenzo's Oil*, which opened in the United States last month and is released in the United Kingdom tomorrow. The film is brilliantly produced and the race for a 'cure' is acted out with extraordinary virtuosity. I haven't experienced such palpable suspense since Truman Capote's *In Cold Blood* appeared serially in *The New Yorker*. In the course of the film, the audience learns about X-linked autosomal recessive inheritance and fatty-acid metabolism in a graphic and comprehensible manner.

Unfortunately, this seductive film comes at a time when the anti-intellectual populism of Haider, LePen, Perot and others of their ilk has created a receptive audience for an anti-establishment diatribe. We learn in the course of two hours that nurses are heartless, physicians pompous fools and parent support groups as mindless as a herd of sheep. Michaela and Augusto Odone to the rescue! They go to the library and learn about fatty-acid metabolism, sponsor a scientific conference and come up with a solution when they

find that the fat-free diet recommended by the insouciant doctors is doing their child no good. First they pump in the triglyceride of oleic acid via a nasogastric tube to Joan Sutherland's rendition of *Casta diva* from *Norma* as background music. And later, when they find the true grail — erucic acid triglyceride from rapeseed oil — the childless, hetero-

symptoms does not seem convincingly to alter the onset of the disease. There is a complex array of phenotypes in ALD and the penetrance of the disease may result from so far unidentified autosomal genes. *Lorenzo's Oil* may be a cinematic triumph, but theatrical truth and scientific truth are not necessarily miscible.

Between the Romanovs and the Odone, concerned parents have often banded together to seek medical solutions for their afflicted children. Frequently they have acted effectively and expeditiously while the usual channels of the funding authorities have dragged their heels. The Muscular Dystrophy Association, which is capable of the



Slippery business — Nick Nolte as Augusto Odone in *Lorenzo's Oil*.

zygous maternal aunt of the propositus swallows a test meal of yet another triglyceride to the strains of *Una furtiva lagrima* from *L'Elisir d'Amore*. By the end of the film, when a chorus of children on the screen tell us they are all taking Lorenzo's oil, and a toll-free number is flashed up, you may be tempted to rush out and sprinkle your next green salad with this concoction. Before you do, beware.

The US Food and Drug Administration, roused by the Odone from its customary torpor, quickly approved Lorenzo's oil. Hugo Moser, one of the authors of the paper on page 726 of this issue, reporting the identification of a putative X-linked ALD gene (see also News and Views on page 682), will soon report that some 40 per cent of patients with ALD develop thrombocytopaenia from this treatment, as the triglyceride disrupts the platelet membrane. Furthermore, feeding the stuff to boys with the ALD defect before the onset of

most flamboyant public relations, supported Dr Louis Kunkel in his successful effort to clone the dystrophin gene before the National Institutes of Health recognized the importance of his work. There are other compelling examples of this sort such as to encourage individual and group initiatives to seek solutions to biomedical problems. Indeed, the paper on page 726 cites such private funding sources in the quest for the ALD gene. A surprising answer to the underlying defect in ALD has been uncovered and it may in the long run provide a strategy for gene therapy in this distressing disorder. Augusto Odone, with his spoor of countless cigarette butts, and his monomaniacal wife, Michaela, have not set a good example for biomedical progress. That is why this entrancing film is, at base, pernicious. □

Fred S. Rosen is at the Center for Blood Research, 800 Huntington Avenue, Boston, Massachusetts 02115, USA.

Tourist attraction

W. F. Bynum

The Scientific Traveler: A Guide to the People, Places and Institutions of Europe. By Charles Tanford and Jacqueline Reynolds. Wiley: 1992. Pp. 335. £10.95, \$17.95 (pbk).

ALAN Bennett recently remarked that he would rather look at a good quality postcard than trudge around museums, scrutinizing the paintings themselves. Some travel books are written for people like him: volumes that can be enjoyed thousands of miles away from the scenes they describe and that are unlikely ever to be packed in the suitcase. Another class of travel guide demands to be taken along; indeed, is rarely valued seriously except *in situ*.

Charles Tanford and Jacqueline Reynolds clearly intend *The Scientific Traveler* to be packed, and no one would get much satisfaction in staying at home and looking at its poor-quality illustrations. The volume is no substitute for the real thing. At the same time, I doubt if it will inspire many to take up the art of scientific travel. Instead, it succeeds on two fronts, providing historical background to the scientific traditions of Europe, and useful information about some of the things of scientific interest to be seen if one finds oneself in Germany, France, Ireland or some other European country.

About two-thirds of the volume is concerned with European science and scientists. Like any guide book, geography determines its structure. Seventeen chapters describe the scientific traditions of as many countries, the combination of Poland and Russia into a single chapter being offset by an introductory chapter on "Our European Heritage". England, Scotland and Ireland merit separate consideration; Wales, Portugal, Norway, Finland and Belgium are omitted. Italy, France, Germany and England rate the longest chapters, with their scientific histories recounted largely through the biographies of some of their principal scientists and, less conspicuously, their main scientific institutions. Although the authors discount their book's pretensions as a history of science, the life stories of some 300 scientists are told in such a way that the various national traditions in science can be appreciated. *The Dictionary of Scientific Biography* seems to be the authors' chief source, but they maintain a reasonable level of accuracy and possess a good eye for anecdote.

After the heritage comes the travel. For each country, a second section on



THE Eocene oil shales of Messel, near Frankfurt, Germany, are remarkable for the unusually complete and detailed picture they give of life some 50 million years ago. Shown here are a fossilized jewel beetle (Buprestidae; left, $\times 4.2$) and a winged 'queen' of a small ant species (Formicidae; $\times 2.2$), both common at the site. The pictures are taken from *Messel* edited by S. Schaal and W. Ziegler, which contains hundreds of colour illustrations of the fossils and descriptions of their ecology, biogeography and evolutionary significance. Oxford University Press, £50.

"Principal Places to Visit" details the tangible legacies of European science and scientists. These range from surviving birthplaces to graves, and include museums, laboratories, statues, plaques and, in several cases, geologically important scenery. Thus we are directed to Neuchâtel in Switzerland to admire the glacial deposits first elucidated by Louis Agassiz, and to Kirbymoorside in Yorkshire, to visit Kirkdale cave, where William Buckland identified the fossilized bones of hyenas and other species extinct in Britain. The museum and cliffs at Lyme Regis recall the fossil-hunting of Mary Anning and the Victorian discovery of *Ichthyosaurus* and other giant vertebrates. Archaeological and prehistoric sites also get good coverage, with descriptions of stone circles in England and Scotland, cave paintings in Lascaux, Tarascon-sur-Ariège and Altamira, and the museum at Abbéville, where Jacques Boucher de Perthes uncovered so many worked flints.

For the most part, however, Tanford and Reynolds guide us to the urban legacies of science. London tops the list with 11 separate listings, such as the Royal Observatory, the Natural History Museum and the Royal Institution. Paris has eight, including La Villette, the Institut Curie and the Pasteur Museum. In Paris, the Jardin des Plantes has a full entry, whereas Kew Gardens in London gets none. Westminster Abbey rates a section, whereas Père Lachaise gets only a passing mention. Most other major

European cities — Rome, Berlin, Edinburgh, Stockholm, Geneva — get reasonable coverage, as quite naturally do Padua, Leiden and Edward Jenner's Berkeley. Out-of-the-way places include the Mendelianum in the monastery in Brno; the island of Ven, where Tycho Brahe's observatory stood; and the Temple of Serapis, near Naples, made famous by the frontispiece of Charles Lyell's *Principles of Geology*.

As the authors make clear, this book represents the personal selection of two indefatigable travellers, who have retired to England after distinguished scientific careers at Duke University. They have seen the places they describe and include brief instructions on how to get there, opening hours and contact telephone numbers. It is as up to date as the political turmoil in Eastern Europe permits and includes, for instance, a description of the newly opened Boerhaave Museum in the old St Caecilia Hospital in Leiden.

Most good general guide books contain some information on places of scientific interest, but I know of no volume in print that focuses on science and which has such broad geographical coverage. I read it from cover to cover without getting itchy feet, but I shall certainly take it along the next time I travel in Europe. □

W. F. Bynum is at the Wellcome Institute for the History of Medicine, 183 Euston Road, London NW1 2BE, UK.

The particle connection

C. H. Llewellyn Smith

Dreams of a Final Theory. By Steven Weinberg. *Pantheon/Hutchinson*: 1993. Pp. 260. \$25, £16.99.

THIS important book will stimulate, provoke and possibly outrage, but never bore, its readers. Steven Weinberg's intention is "to lay out the issues raised by the idea of a final theory as part of the intellectual history of our times, for readers with no prior knowledge of physics or higher mathematics". This suggests that the book is simply a popularization for nonscientists. In fact, it is an exceptionally clear and profound exposition of the views of a leading scientist on issues that are by no means uncontroversial, such as reductionism, the role of beauty and mathematics in the formulation of physical laws, the nature of scientific explanation, the status of 'emergent' phenomena, the (lack of) utility of philosophies of science, positivism, the possibility of discovering the handiwork of God in nature, and the possible existence and nature of a final theory, which Weinberg thinks exists and may not be far from our reach.

Many working scientists will be delighted to read Weinberg's carefully marshalled arguments supporting points of view that many of us hold but have never been forced to articulate. His explanation of the interplay of theoretical reasoning and experiment, which he illustrates with the discoveries of relativity, quantum electrodynamics and the 'standard model' of elementary particles, is outstanding and will strike a chord in anyone who has been engaged in fundamental research. More controversially, Weinberg argues — convincingly in my opinion — that "The reason we give the impression that we think that elementary particle physics is more fundamental than other branches of physics is because it is", although he is careful to say that this "does not mean that it is more mathematically profound or more needed for progress in other fields or anything else but that it is closer to the point of convergence of all our arrows of explanation". The reader who disagrees should read the book, as should those who believe that complex (chaotic) systems exhibit new kinds of fundamental laws, a view that Weinberg convincingly demolishes. Less controversially, among scientists at least, he also does a wonderful demolition job on the view that scientific knowledge is largely subjective.

The one point of physics on which I

am unable to follow Weinberg is his espousal of Everett's many-worlds interpretation of quantum mechanics, which seems to me not only extravagant but ill defined, and his belief that "The one part of today's physics that seems likely to survive unchanged in a final theory is quantum mechanics", not only because it works so well but because "no-one has been able to think of any way to change quantum mechanics in any way that would preserve its successes without leading to logical absurdities". I remember hearing another Nobel prizewinner use the same argument to 'prove' that a 'renormalizable' theory of weak interactions could not exist, at about the time that Weinberg constructed one, for which he won his Nobel prize! Weinberg may be right, but I would not be surprised if our understanding of quantum mechanics changes radically in the future, and (like Newton's laws) it may turn out to be only approximate; in either case, the construction of a final theory would presumably be impossible until this occurs.

Weinberg states that this "is not a book about the [Superconducting] Super Collider [SSC]," although the debate over the project forced him "to try to explain what we are trying to accomplish in our studies of elementary particles". But the case for the SSC runs through the book and provides the climax. This is a pity, because it reads like special pleading, and is likely to prove more ephemeral than the rest of the book.

Weinberg begins the book by stating that "The century now coming to a close has seen in physics a dazzling expansion of the frontiers of scientific knowledge... But now we are stuck" and that the SSC was planned "in order to break out of this impasse". Fair enough, and later he explains brilliantly the frustration produced by the incredible success of the standard model of quarks and leptons governed by electroweak and strong forces, which currently describes all data with ever improving accuracy while becoming more and more obviously incomplete. However, I consider that Weinberg overplays his hand in asserting the absolute necessity, and in implying the possible sufficiency, of the SSC in the search for a final theory.

Progress to a better theory "beyond the Standard Model" is blocked in particular by ignorance of the mechanism that breaks the symmetry between electromagnetic and weak forces. Weinberg states that "The only sure way to settle this question is to do experiments in which a trillion volts is made available for the creation either of Higgs particles or of massive particles held together by extra-strong forces. For this purpose it turns out to be necessary to give a pair of colliding protons a total energy of

about 40 trillion volts [the energy of the SSC]". It is certainly true that the SSC will explore an energy range in which it is almost certain that new phenomena associated with electroweak symmetry breaking will appear. But the elusive Higgs particle may well be discovered in experiments at the Large Electron-Positron Collider (LEP) at CERN, the European Laboratory of Particle Physics, long before the SSC is built. On the other hand, if there is in nature an underlying 'super symmetry' (connecting particles obeying Fermi and Bose statistics), as Weinberg believes, it may not be possible to detect Higgs particles produced by the SSC (although there would be additional 'super particles' awaiting discovery by LEP and the SSC).

Furthermore, if the question of electroweak symmetry breaking can be settled by the SSC, it can almost certainly also be settled by the European version — the Large Hadron Collider (LHC) — proposed at CERN. Dismissing the LHC in one paragraph, Weinberg states that it "would cost much less than the Super Collider" but its energy "would be limited to less than half of the Super Collider". This is somewhat misleading, because the physics potential does not increase linearly with the energy and also depends on the luminosity, which is maximized in the design of the LHC (but not in the current SSC design), thereby giving the LHC the same potential to produce heavy particles as the SSC, although the experiments will be harder. (Furthermore, the LHC can be used to collide nuclei and also collide electrons with protons at unprecedented energies.)

Although Weinberg is careful to say that beyond the questions that he expects to be answered by the SSC "there is a level of deeper questions... that cannot be directly addressed by any accelerator now conceivable", he nevertheless raises the prospect in readers' minds that the final theory may be discovered in their own lifetimes; and the statement, with which the final paragraph begins, that "No-one can say whether any one accelerator will let us make the last step to a final theory", certainly seems to imply that the SSC might do the job. Although I find this line of promotion disturbing, Weinberg makes an excellent case for the SSC (and also by implication for the much less costly LHC) and I have dwelt on his arguments only because this part of the book has attracted considerable attention. The case he makes for a realist, reductionist approach to physics is even better and is likely to be read for many years to come. □

C. H. Llewellyn Smith is in the Department of Physics, University of Oxford, OX1 3NP, UK.

Avian conflicts

P. L. Schwagmeyer

Dunnoch Behaviour and Social Evolution. By N. B. Davies. *Oxford University Press*: 1992. Pp. 272. £35, \$70 (hbk); £13.50, \$28 (pbk).

MUCH of behavioural ecology focuses on how individuals optimize their reproductive success under constraints imposed by phylogeny and the environment, the latter including resources, predators, parasites and members of the same species. Conspecifics may assume the role of ally, foe or a bit of both. That these roles ultimately depend on genetic interests is illustrated beautifully by Davies' studies of dunnoch reproductive behaviour.

Although bland in appearance, dunnocks defy even sixties' norms of sexual behaviour. While they often practice 'conventional' avian monogamy, they also sometimes form trios (two unrelated males plus a female or vice versa), and sometimes neighbouring females are jointly defended by two or more males. Contrary to traditional views of how avian mating associations form, Davies argues that dunnoch females essentially ignore males while establishing their breeding territories, and males then compete to defend resident females (that is, like typical mammals). Which mating system occurs depends on how easily female territories can be monopolized by males: the larger a female's territory, the more likely she is to have two or more males try to defend her.

The shared defence of a mate creates plenty of conflicts of interest. Although two males who simultaneously attempt to defend a female eventually reach a dominance-based truce, the onset of breeding renews bickering. And with apparent good reason, for if dominant males mateguard effectively, they can sire all the offspring, but if they lapse, their share declines. DNA fingerprinting reveals that a male's paternity is related to the relative time he gains sexual access to a 'shared' female; on average, the dominant male sires about 55 per cent of the brood.

Genetic fingerprinting also shows that each male's participation in feeding the young hinges on gaining sexual access to the female. And therein lies a key insight for understanding why females actively resist being monopolized while fertile: a female that mates with both the dominant and subordinate males acquires the assistance of both in rearing her offspring, while avoiding sabotage of the clutch by the subordinate male. Davies shows that females in polyandrous trios more than double their production of fledged young if they mate

with both males, providing one of the best-documented advantages to vertebrate females of mating with several different males.

The divergence of male and female mating interests and the ups and downs of male-male interactions are just two of several conflicts of interest discussed by Davies. Among the others are conflicts over parental effort and the adversarial relationship between dunnocks and cuckoos. The organization and integration of these topics makes the book useful even to those familiar with his work. It should be particularly instructive for students, because the work addresses questions typical of the general field of behavioural ecology and demonstrates the value of combining experiment and description. And all from the study of mousey little birds breeding a short bicycle ride from Davies' office. □

P. L. Schwagmeyer is in the Department of Zoology, University of Oklahoma, Norman, Oklahoma 73019-0235, USA.

Joint account

D. A. Isenberg

All About Arthritis: Past, Present, Future. By Derrick Brewerton. *Harvard University Press*: 1992. Pp. 317. \$29.95, £15.95.

DERRICK Brewerton, until recently a professor of rheumatology at the University of London, has written an account for the general reader about the nature and causes of arthritis, and the historic search for a cure. The diversity of rheumatic conditions does not make this an easy task, and his narrative is invariably fragmented. Nonetheless, he skilfully guides us through the complexities of bacteria, viruses and host defence; the nervous system; the structure and function of joints; the mosaic of factors (hormonal, genetic and environmental) involved in arthritis; and the prospects for new forms of therapy. He is particularly elegant when dealing with the complex (and controversial) influence of personality and emotion, perhaps best illustrated by the current epidemic of patients diagnosed as having fibromyalgia and myalgic encephalomyelitis. Throughout, the author goes to great lengths to emphasize the global nature of arthritis.

One of the book's main strengths lies in the lucid interweaving of historical vignettes on the founding fathers (and occasionally mothers) of rheumatology and inflammation. Some, such as Gregor Mendel, Louis Pasteur, Robert Koch and Paul Ehrlich, are well known; but

equally stimulating tales about lesser-known individuals abound, such as John Lawrence, who set out in the north of England with a mobile unit complete with clinical laboratory and X-ray facilities to blaze a trail for the epidemiology of rheumatology, and Pamela Bjorkman, who dedicated eight years to determining the structure of the antigen HLA-A2. Brewerton also clearly enjoys retelling stories of those scientists whose important contributions were little appreciated or even ridiculed in their own lifetime, such as Rosalind Franklin (whose X-ray crystallographic work established virus structure) and Martinus Beijerinck (who established the modern concept of viruses). Even Jean-Martin Charcot was much derided for his suggestion that "nerves might nourish the bones and joints so that, in disease, joint failure [is] due to a lack of a trophic factor". But today the French neurologist deserves recognition as the first person to speculate about the existence of neuropeptides.

Although the book is often sparkling reading, the bobbing and weaving between clinical descriptions and basic science is likely to irritate. And despite the useful glossary of terms, I suspect that the author has overestimated the basic knowledge of his intended audience. There are also some surprising omissions. In the discussion of the causes of systemic sclerosis, the terrible story of the link between this disease and contaminated cooking oil in Spain surely merits more than a mention. And there is nothing about diet as a cause of arthritis. Although this area is notorious for some dubious claims, good work (for example, implicating dietary fat in lupus development) has been published and deserves wider recognition.

It is odd, especially given the author's close interest in the subject, that he ignores the long controversy about whether *Klebsiella* infection is involved in the development of ankylosing spondylitis. More recently, Alan Ebringer, who first suggested the link, has put forward the even more controversial suggestion that *Proteus* infection might be important in rheumatoid arthritis, although others now argue strongly that a slow-growing form of mycobacterium is involved. Here I would have liked Brewerton's views — after all, as Ehrlich was apparently fond of remarking, "what matters in research is not the initial discovery but what one thinks and does next". These deficiencies, however, do little to detract from the many gems in this generally well polished book. □

D. A. Isenberg is in the Bloomsbury Rheumatology Unit, Arthur Stanley House, 40–50 Tottenham Street, London W1P 9PE, UK.

Effects of an endothermic phase transition at 670 km depth in a spherical model of convection in the Earth's mantle

Paul J. Tackley*, David J. Stevenson†, Gary A. Glatzmaier‡ & Gerald Schubert§

* Seismological Laboratory, California Institute of Technology, Pasadena, California 91125, USA

† Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, California 91125, USA

‡ Earth and Environmental Sciences Division and Institute of Geophysics and Planetary Physics, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA

§ Department of Earth and Space Sciences and Institute of Geophysics and Planetary Physics, University of California, Los Angeles, California 90024, USA

Numerical modelling of mantle convection in a spherical shell with an endothermic phase change at 670 km depth reveals an inherently three-dimensional flow pattern, containing cylindrical plumes and linear sheets which behave differently in their ability to penetrate the phase change. The dynamics are dominated by accumulation of downwelling cold material above 670 km depth, resulting in frequent avalanches of upper-mantle material into the lower mantle. This process generates long-wavelength lateral heterogeneity, helping to resolve the contradiction between seismic tomographic observations and expectations from mantle convection simulations.

THE question of whether an endothermic phase transition associated with the seismic discontinuity at 670 km depth could enforce layered convection in the Earth's mantle has received much attention^{1–5}, and has profound implications for the Earth's thermal and chemical structure and evolution⁶. Early two-dimensional numerical modelling¹ seemed to show that an unrealistically large Clapeyron slope of -6 MPa K^{-1} would be required for layering to occur. Recent work in two-dimensional cartesian^{2,3,48} and spherical axisymmetric^{4,5} geometries with realistic phase-change parameters has, however, indicated significant layering, as well as complex new phenomena and modes of time-dependence in the flow.

To understand the Earth it is essential to determine how the effects observed in these two-dimensional studies are modified in three-dimensional geometry. Here we present results from a numerical simulation of fully three-dimensional compressible mantle convection in a spherical shell with an endothermic phase change at 670 km depth. The spatial resolution and Rayleigh number are much higher than in previous spherical models with no phase change^{7–9}. Although there are some similarities with the two-dimensional results, the simulated flow pattern is inherently three-dimensional, with features that penetrate the phase change being exclusively cylindrical in form. The upper mantle is characterized by interconnected linear downwellings which do not penetrate. At the intersections of these downwelling sheets, cold material accumulates above 670 km depth, building up until huge catastrophic breakthroughs^{1,4,5} into the lower mantle are precipitated, flushing the local upper mantle contents through broad cylindrical downwellings to the core-mantle boundary (CMB). These events occur in a globally asynchronous manner, with typically three to four in progress at any particular time. This process generates large-amplitude long-wavelength lateral heterogeneity, which may go some way towards reconciling observations of long-wavelength dominance from seismic tomography^{10–15} with the much broader spectrum predicted by high-Rayleigh-number convection simulations^{16,17}.

The behaviour of downwellings is similar to that observed in a completely basally heated three-dimensional cartesian layer with two phase changes¹⁸. We find however that the use of spherical geometry and realistic internal heating greatly modifies the heat flow characteristics, geometry of lower-mantle flow,

upwelling plume structure and global time-dependence, as well as aiding direct comparison with seismic tomography.

Model characteristics

Our model incorporates the effect of the endothermic transition from spinel to perovskite plus magnesiowüstite occurring at $\sim 670 \text{ km}$ depth, consistent with a peridotitic mantle composition¹⁹. Some authors have included additional phase changes in their models^{2,3,5,18,48}. These slightly alter the propensity towards layering and cause some other second-order effects, but considering the other approximations and uncertainties in current mantle modelling we choose to restrict these calculations to an examination of the first-order effect due to the endothermic phase change at 670 km depth. Our preliminary modelling results with the 400 km phase change added display little qualitative difference in behaviour.

The equations of compressible self-gravitating flow are solved and integrated in time using a spectral-transform method¹⁷, implemented on the Intel Touchstone Delta parallel supercomputer system at the California Institute of Technology. As the mantle is characterized by very high Prandtl number, inertial terms are ignored. The anelastic approximation is used, eliminating acoustic waves which propagate many orders of magnitude faster than convective velocities. Entropy, pressure, gravitational potential and poloidal mass flux potential are expanded laterally in spherical harmonics, up to degree and order 127, and the nonlinear products associated with advection and viscous dissipation are evaluated in grid space using 384 longitudinal points by 192 latitudinal points. Variables are expanded vertically in separate Chebyshev series for the upper and lower mantles, with 33 radial levels in the lower and 17 radial levels in the upper mantle. Use of two Chebyshev expansions matched at 670 km depth gives good vertical resolution at the CMB, at 670 km depth and at the surface, where it is most needed. The fields are adequately resolved because the horizontal and vertical variances of variables fall by many orders of magnitude between maximum values and truncation^{7–9,17}.

As the phase boundary deflection observed in the Earth²⁰ is at least an order of magnitude smaller than the characteristic size of convective features in our model, it is not necessary (to a good approximation) to resolve the exact details of the phase change. Thus, deflection of the phase change resulting from



FIG. 1 Cold downwellings for final frame of simulation. Blue surface is an isocontour showing where the temperature is 110 K lower than the horizontally averaged value. Green surface is the core. A network of interconnected linear downwellings is visible in the upper mantle, with three huge cylindrical downwellings in the lower mantle, spreading out into pools of cold material above the core-mantle boundary.

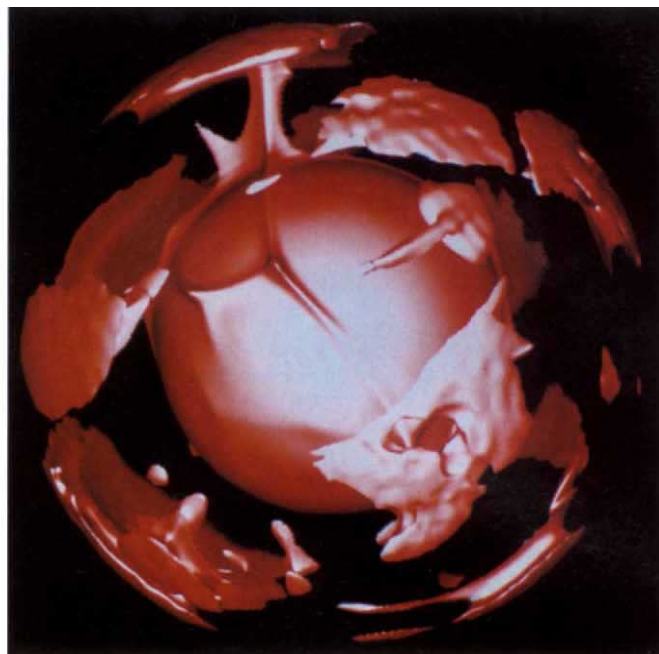


FIG. 2 Hot upwellings for final frame of simulation. Red surface is an isocontour of superadiabatic temperature, showing where the temperature is 110 K higher than the reference-state adiabat. A single plume from the core-mantle boundary feeds a hot region in the upper mantle. Note that most broad hot regions in the upper mantle are not directly linked to lower-mantle structures.

lateral temperature variations at 670 km depth is represented as a sheet mass anomaly at this depth, resulting in discontinuous normal stress between the two Chebyshev regimes. This approach has the advantage of a phase loop (the pressure or depth interval over which the multivariant phase change occurs) with zero width—important because the phase loop width affects the inhibition of flow across the boundary⁵, and recent experiments¹⁹ constrain this width to be a few kilometres at most. For numerical reasons, the latent heat release (absorption) that accompanies upward (downward) motion through the phase change must be spread out 25 km on either side of the interface. This treatment of the phase change has been validated by computing two-dimensional cartesian and spherical axisymmetric results and comparing them to previously published results^{1,4}.

Parameters. Entropy, pressure and gravitational potential are expanded as perturbations relative to a self-gravitating adiabatic reference state. The Murnaghan equation is assumed, leading to a polytrope¹⁷ (that is $P \propto [\rho^{1+1/n} - \rho_0^{1+1/n}]$, where n is the polytropic index). A constant Grüneisen parameter of 1.0 and polytropic index of 0.5 gives a reasonable fit to the Earth model PREM²¹. Implicit in this treatment is the depth-dependence of material properties such as density bulk modulus and thermal expansivity, with the latter varying from 3.0×10^{-5} at the surface to 2.2×10^{-5} at 670 km depth and 1.2×10^{-5} at the CMB. Heat capacity is assumed constant at $1,250 \text{ J kg}^{-1} \text{ K}^{-1}$. Thermal conductivity and viscosity are specified as functions of depth only, with dynamic viscosity increasing roughly exponentially from $1.7 \times 10^{22} \text{ Pa s}$ at the surface to $1.9 \times 10^{22} \text{ Pa s}$ at 670 km depth and $2.1 \times 10^{23} \text{ Pa s}$ at the CMB. Thermal conductivity is $2.3 \text{ W m}^{-1} \text{ K}^{-1}$ at the surface, increasing with depth as the fourth power of density, giving a lower-mantle increase consistent with experiments²² and theory²³. The boundary conditions at the surface and CMB are stress-free and isothermal, with the temperatures at the CMB and surface fixed at 3,450 K and 1,060 K respectively. Of the total temperature drop across the mantle, 1,250 K is superadiabatic and 1,140 K is due to adiabatic compression. The unrealistically high surface temperature is a consequence of treating the viscosity as a function of depth only

rather than a function of temperature, so rheologically stiff plates are not produced. Absolute temperatures in such a calculation are not meaningful, however, except to the extent that they are adjusted to be realistic at some depth (for example, the upper mantle); only temperature differences are relevant to the convective style and vigour.

Because the ratio of internal to basal heating strongly affects the degree of layering in two-dimensional models (W. R. Peltier, personal communication), it is important to match the internal heating rate and core heat flux of the Earth as closely as possible. We therefore use an internal heating rate of $2.75 \times 10^{-12} \text{ W kg}^{-1}$, compatible with chondritic values. The volume averaged Rayleigh numbers resulting from internal heating and super-adiabaticity⁹ are 1.8×10^7 and 1.2×10^6 respectively, an order of magnitude higher than in previous studies⁷, but almost an order of magnitude less than those characterizing the Earth. Realistic Rayleigh numbers are obtainable in two-dimensional calculations^{2,3,5}, which are useful for predicting how the effects observed here might scale to the Earth-like regime.

We take the value of the Clapeyron slope to be -4 MPa K^{-1} , the preferred value from recent experimental results²⁴, although at the high end of the range from previous experiments¹⁹. We choose a high value so that we can determine and characterize the maximum effect that the phase change could have on the flow. Although this value has been observed to cause very strong layering in one axisymmetric result⁴, the inhibiting effect of the phase change in that simulation was greatly enhanced by the use of a much lower thermal expansivity at 670 km depth than is consistent with experimental and theoretical estimates²⁵, and by the imposed two-dimensionality.

The simulation was started from a case with no phase change and run for about 15,000 timesteps, corresponding to about 3 billion years; the results we present here are characteristic of the last 9,000 timesteps, after the system has overcome the initial transient adjustment to the presence of the phase change.

Observables

We obtain a mean surface heat flow of $2 \times 10^{13} \text{ W}$, of which

~40% comes from the core. The total is similar to that of the Earth, but the basal heatflow is considerably larger than most estimates for the Earth^{26–28}. Even so, it drives very little plume activity, as discussed below. It is possible that the mantle heating rate should be substantially augmented by the effect of secular cooling^{26,27,29} which is missing from the model because the heating is treated as time-independent, but this is an issue for the future.

Maximum convective velocities are $\sim 40 \text{ mm yr}^{-1}$ and average surface velocities typically 6 mm yr^{-1} —an order of magnitude lower than plate velocities, although a direct comparison may be misleading because our simulation lacks rigid plates. The principal difference between our parameters and the Earth probably lies in the use of higher than realistic viscosities, which is necessitated by computing limitations. At the higher Rayleigh number that would result from lower viscosities, the phase change has a stronger inhibitive effect on the flow^{1,3}; thus our calculation may underpredict the degree of layering that would occur in the Earth.

Flow pattern. In Fig. 1, which illustrates cold features for the final frame of the calculation, a network of interconnected cold downwelling sheets, which do not penetrate the phase change, is observed in the upper mantle. The distance between sheets that are roughly parallel is typically 3,000–8,000 km, a scale which is consistent with present subduction zones on the Earth. They have some small-scale complexity due to local boundary-layer instabilities, which would probably be suppressed by the high-viscosity lithospheric plates on the Earth. At the intersections of these sheets, large pools of cold material form above the 670 km phase change. This cold material is gravitationally unstable and after building up sufficiently triggers a sudden avalanche into the lower mantle, in the form of a large-diameter ($\sim 1,000 \text{ km}$) cylindrical downwelling plume. This downwelling acts as a conduit to the CMB, effectively emptying the cold material from the local upper mantle to a large pool at the base of the mantle, despite the increase in viscosity with depth. The downwelling then shuts off completely and does not recur in exactly the same place during the simulation, although many such events may occur in the same general area, and a total of

about 15 events are observed during this part of the simulation. Thus, these events cool all areas of the lower mantle and core. At any one time, several of these flushing events are in progress at different places around the sphere, triggering in a globally asynchronous manner. Owing to the spherical geometry, the surface area of the CMB is $\sim 35\%$ of the surface area of the 670-km interface, and thus the combined effect of flushing events occurring in different places around the sphere is to surround the core with cold material, resulting in a heat flow (40% of surface heat flow) considerably higher than most estimates for the Earth^{26–29}, but with very little upwelling plume activity.

Figure 2 shows the corresponding hot, upwelling regions for the final frame. The most prominent features are the broad hot regions in the upper mantle, which are generally not associated with any deep features in the lower mantle. The lateral heterogeneity is much greater in the upper mantle than in the lower mantle, as shown quantitatively below. At the CMB, ridges of hot material are observed. These are swept around by the enormous injections of cold upper mantle material caused by flushing events, and occasionally a short-lived, transient plume is formed at the intersection of these ridges, rising to the 670-km interface and injecting hot material into the upper mantle. One of these can be observed in Fig. 2. Thus, in both directions, cylindrical forms (plumes) are seen to penetrate the 670-km phase change, whereas linear forms (sheets) do not. The upward flow in the upper mantle and in the top of the lower mantle is generally the weak, distributed return flow characterizing mainly internally heated convection^{7–9,30,31}.

The radial dependence of the flow structure is clearly visible in Fig. 3a and b, which shows the superadiabatic temperature field in two cross-sectional slices for the final frame of the simulation. The upper mantle is extremely heterogeneous on long wavelengths, containing some broad (up to $\sim 10,000 \text{ km}$) regions of anomalously cold or hot material, as well as regions with the classical small aspect-ratio convection cells^{32,33}. In the lower mantle, two large cylindrical downwellings are visible (Fig. 3b), and a thick layer of cold material surrounds the core. One broad upwelling plume from the CMB (Fig. 3a) penetrates the 670-km phase change and injects hot material into the upper

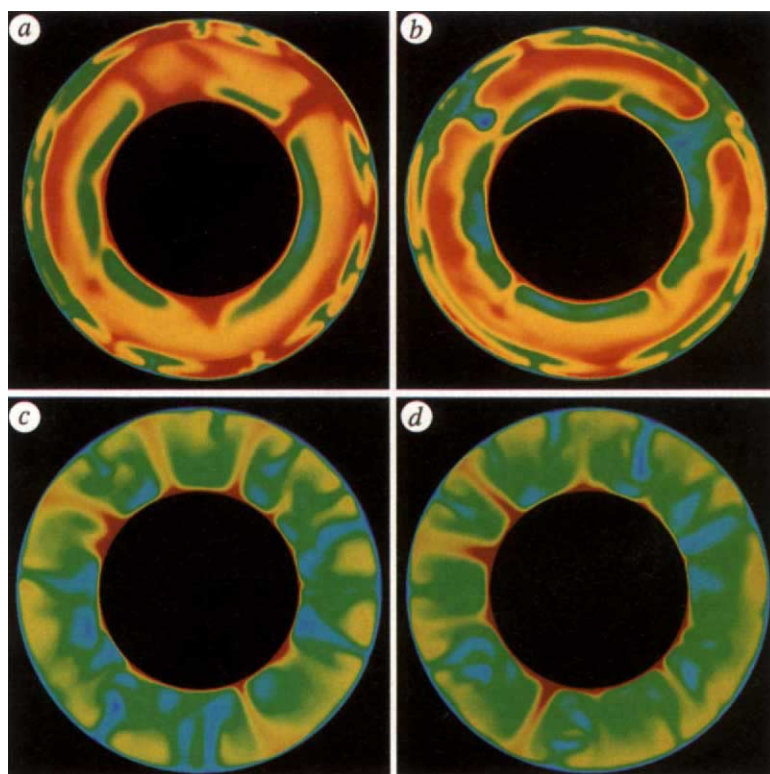


FIG. 3 Superadiabatic temperature field on different cross-sectional slices. *a* and *b*, The final frame of the phase change simulation. Scale ranges from $-1,050 \text{ K}$ to $+350 \text{ K}$. *c* and *d*, A typical case with no phase change. This is representative of mainly internally heated whole-mantle models^{7–9,30,31}, with Rayleigh numbers for internal and basal heating of 1.4×10^7 and 5.5×10^5 respectively, a factor of 10 viscosity increase with depth, and a basal heat flow of $\sim 16\%$ of the surface heat flow. Scale ranges from -780 K to $+220 \text{ K}$.

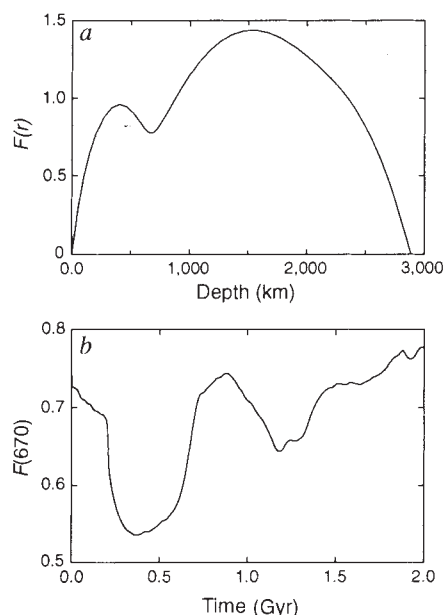


FIG. 4 *a*, Radial mass flux diagnostic $F(r)$ for the final frame of simulation. *b*, Time series of $F(670)$ for the last 9,000 timesteps (2 Gyr). $F(r)$, defined in ref. 5, is the spherically averaged modulus of radial mass flux, normalized so that the integral over non-dimensional depth is unity.

mantle, but this is the only such feature in the whole lower mantle. The inclusion of realistic temperature-dependent viscosity would result in a low-viscosity layer immediately above the CMB because of the thermal boundary layer, which may result in more upwelling plumes. The lack of strong plume activity in our model is not, however, of concern as plumes on the Earth are merely secondary features, carrying only 10–15% of the total heat flux²⁸.

The ability of the cold downwelling lower mantle plumes to reach the CMB, despite the increasing viscosity and decreasing thermal expansivity with depth, is in marked contrast to a similar case with no phase change illustrated in Fig. 3c and d, in which the somewhat smaller downwellings slow down and broaden in the lower half of the mantle, allowing the formation of many upwelling plumes, despite the low basal heat flux. It is possible that a larger viscosity increase with depth²⁸ may slow down the giant drips observed in the phase-change case. They are such large features, however, with so large a reservoir of feeding material, that it seems unlikely that they will be prevented from reaching the CMB.

Radial mass flux. A useful indicator is the radial mass flux diagnostic⁵, defined as the spherically averaged absolute radial mass flux, normalized so that the integral over non-dimensional depth is unity. This diagnostic for the final frame is shown in Fig. 4, together with the time-series of mass flux through the 670-km discontinuity for the last 9,000 timesteps (~2 Gyr) of the simulation. The phase change is observed to have a marked inhibitive effect on the flow, indicated by the minimum in the radial mass flux at 670 km. Examination of this diagnostic for individual spherical harmonic degrees indicates that long wavelengths of flow are virtually unaffected by the phase change, whereas short wavelengths are increasingly inhibited; for spherical harmonic degrees above ~40, flow is effectively confined to the upper mantle.

From the radial mass flux across the interface at 670 km depth it is possible to calculate a 'mixing time', defined as the time required for a mass equal to the mass of the mantle to pass through 670 km. For this case the time is 4.5 Gyr. To scale to the Earth's Rayleigh number (Ra), it is likely that the mass flux

(of downwelling cold material) scales roughly as the heat flux, suggesting a $Ra^{1/3}$ scaling³⁴ which would reduce the mixing time to ~2.1 Gyr for an order of magnitude increase in Ra. This is less than half the age of the Earth, but the increased effect of the phase change at higher $Ra^{1,3}$ may increase this time.

There is some time dependence of the flow through the phase change at 670 km depth, with the mass flux diagnostic at 670 km depth ($F(670)$) varying between 0.5 and 0.8 during this simulation. This time dependence is much weaker than that observed in two-dimensional geometry^{4,5,48}. Owing to the large effective aspect ratio in our spherical simulation, several flushing events are typically occurring at any given time, so the convection is never strongly layered on a global scale. As a result globally averaged diagnostics such as $F(670)$, mean temperature, and so on are not greatly affected by an individual flushing event. Given the observed preference for a penetration in cylindrical rather than linear forms, it is also possible that in two-dimensional geometry (where up- or downwellings are restricted to being linear), cold material builds up to a greater extent before it can be flushed into the lower mantle, and thus these events are more violent. It is possible that our flushing events would be more abrupt and violent if we could decrease the viscosity to Earth-like values, as observed in two-dimensional models (ref. 48, and W. R. Peltier, personal communication). We believe however, that with an Earth-like viscosity, flushing events would still overlap in time, and the global time dependence would be intrinsically weaker in full three-dimensional spherical geometry than in spherical axisymmetric or two-dimensional cartesian geometries.

Spherically averaged temperature. The horizontally averaged ($l=0$) temperature field is shown in Fig. 5. A temperature drop of several hundred degrees occurs around 670 km depth. In addition a small kink is observed at 670 km, due to the release or absorption of latent heat by material advected across the phase change. The net thermal gradient at 670 km depth is low, and thus the conductive heat flow across the 670-km interface is only ~10% of the surface heat flow, with most of the heat flux across 670 km depth being advected. On local scales, much steeper temperature gradients are observed in regions far from cylindrical downwellings and most of the radial heat flux may be conductive.

Lateral heterogeneity and seismic tomography

The effects of the phase change on both the horizontal spectrum of density anomalies and the depth-dependence of total power⁵ is pronounced, resulting in a dominance of long-wavelength power in the upper mantle and base of the lower mantle, and a concentration of total power in the same areas (Fig. 6a and b), compatible with seismic tomography^{13,15}. In the upper mantle the dominant frequencies are spherical harmonic degrees $l=2-7$, with a peak at $l=6$, whereas in the deepest ~400 km of the lower mantle, $l=2$ and 3 dominate, with a secondary peak

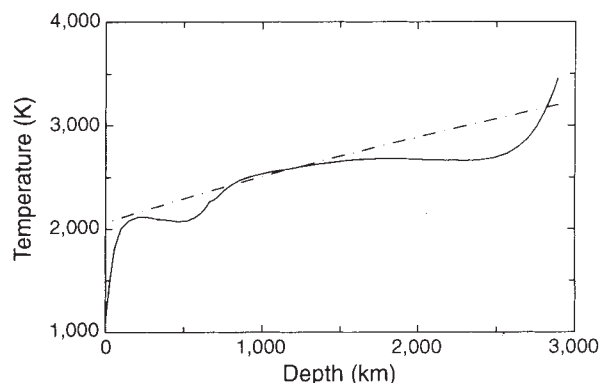
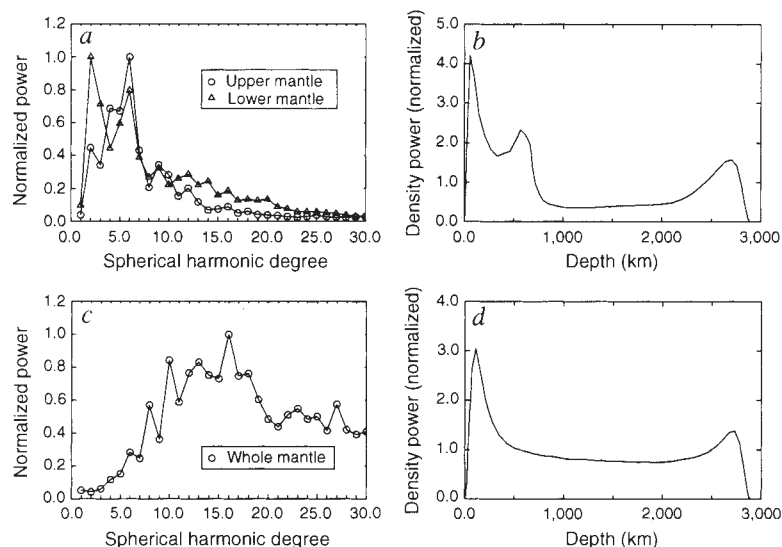


FIG. 5 Solid line, radial profile of spherically averaged temperature for final frame. Dashed line, reference state adiabat.

FIG. 6 *a*, Power in each spherical harmonic degree, vertically averaged through upper and lower mantles, for the final frame. This lateral power spectrum has similarities with that observed by Inoue *et al.* using seismic tomography (plotted in ref. 13, Fig. 5) *b*, Radial profile of total lateral heterogeneity (in density). A peak is observed at 670 km depth, as with two-dimensional results⁵. *c*, Power in each spherical harmonic degree, vertically averaged through the whole mantle, for the case with no phase change (illustrated in Fig. 3c and *d*). *d*, As *b* for the case with no phase change.



at $l = 5-6$. This is very different from the illustrated result with no phase change (Fig. 6c and *d*), in which a broad peak is observed in the range $l = 8-20$; and also with basally heated model at $Ra = 10^6$ (ref. 17; Fig. 1) which peaks at $l = 10$. Indeed, the lateral spectrum is fairly consistent with the observations of long-wavelength dominance from seismic tomography^{10-13,15,35}, particularly that of Inoue *et al.*³⁵ (plotted in refs 13 and 15). The flow lacks the dominant $l = 2$ component found in some seismic models¹⁰⁻¹²; this is probably due to the absence of tectonic plates and continents, as discussed later. The asymptotic fall-off in power with spherical harmonic degree varies with depth, but is roughly proportional to l^{-2} .

Even when filtered to seismic tomographic resolution (for example $l = 10$, Chebyshev degree $n = 13$), the effect of the phase change is still clearly discernible, both in the entropy field and in radial correlations between layers⁴⁹. However, all plume-like upwellings on the lower mantle disappear completely. The strongest of the broad cylindrical downwellings are visible as smeared out cold anomalies, and thus may be discernible in tomographic models. The peak in the radial heterogeneity profile (Fig. 6b) is still visible at $n = 13$, but disappears at $n = 6$. Although the lateral heterogeneity spectrum is dominated by long wavelengths, we need to expand to at least $l \approx 31$ to see all the main features.

Slab Penetration

The issue of whether subducting lithospheric slabs penetrate the lower mantle or are deflected by the 670-km discontinuity is central to the question of mantle layering³⁶⁻⁴². The closest analogues to slabs in our model are the linear downwellings in the upper mantle, and our model results suggest that slabs do not immediately penetrate, but instead build up in the transition zone until enough cold mass has accumulated to precipitate an avalanche into the lower mantle⁴³. This picture would be consistent with (1) strong correlations between past locations of subduction and seismically fast ('cold') regions at the base of the upper mantle and top of the lower mantle⁴⁴, (2) recent tomography of slabs³⁹⁻⁴¹ showing that at least in some areas slabs flatten out along the 670-km discontinuity and appear to stagnate in the transition zone, (3) recent global tomography¹⁴ showing broad seismically fast areas at 600 km depth in places where subducted slabs would be expected to accumulate, and (4) global mapping of topography on the 670-km discontinuity²⁰ favouring models in which subducting slabs are deflected horizontally at the discontinuity.

Differences from the Earth?

Although our model does not include the effects of variable

viscosity, tectonic plates or continents, many of the chief characteristics of the Earth are reproduced, such as the predominance of linear downwellings in the upper mantle (analogous to slabs), wavelength of upper-mantle features (compatible with typical plate sizes), the secondary nature of upwelling plumes, and the similarity of lateral heterogeneity spectra at different depths with seismic tomographic models^{13,15}.

One difference is that junctions of linear downwellings, at which the large pools of cold material build up in our model, are not observed in the Earth except at the intersection of the Pacific, Philippine and Eurasia plates. This may make catastrophic breakthrough more difficult and more violent, as in two-dimensional models^{1,4,43,48}. In addition, the cold slab material is highly viscous and if deflected by the phase change, may have a greater tendency to stagnate in the transition zone.

The addition of rigid plates would probably reinforce the long-wavelength nature of the flow^{45,46}. The plate-tectonic cycle, including subduction-zone locations, may be controlled by the assembly and breakup of supercontinents⁴⁷, which would pump power into the very lowest spherical harmonic degrees⁴⁴. □

Received 30 November 1992; accepted 29 January 1993.

- Christensen, U. R. & Yuen, D. A. *J. geophys. Res.* **90**, 10291-10300 (1985).
- Liu, M., Yuen, D. A., Zhao, W. & Honda, S. *Science* **252**, 1836-1839 (1991).
- Zhao, W., Yuen, D. A. & Honda, S. *Phys. Earth planet. Inter.* **72**, 185-210 (1992).
- Machetel, P. & Weber, P. *Nature* **350**, 55-57 (1991).
- Peltier, W. R. & Solheim, L. P. *Geophys. Res. Lett.* **19**, 432-434 (1992).
- Wyllie, P. *J. Rev. Geophys.* **26**, 370-404 (1988).
- Glatzmaier, G. A., Schubert, G. & Bercovici, D. *Nature* **347**, 274-277 (1990).
- Bercovici, D., Schubert, G. & Glatzmaier, G. A. *Science* **244**, 950-955 (1989).
- Schubert, G., Bercovici, D. & Glatzmaier, G. A. *J. geophys. Res.* **95**, 14105-14129 (1990).
- Nakanishi, I. & Anderson, D. L. *J. geophys. Res.* **88**, 10267-10283 (1983).
- Tanimoto, T. *J. Phys. Earth* **38**, 493-509 (1990).
- Tanimoto, T. *Geophys. J. Int.* **100**, 327-336 (1990).
- Su, W. & Dziewonski, A. M. *Nature* **352**, 121-126 (1991).
- Su, W., Woodward, R. L. & Dziewonski, A. M. *Nature* **360**, 149-152 (1992).
- Su, W.-J. & Dziewonski, A. M. *Phys. Earth planet. Inter.* **74**, 29-54 (1992).
- Jarvis, G. T. & Peltier, W. R. in *Mantle Convection: Plate Tectonics and Global Dynamics* (ed. Peltier, W. R.) 479-594 (Gordon and Breach, New York, 1989).
- Glatzmaier, G. A. *Geophys. Astrophys. Fluid Dyn.* **43**, 223-264 (1988).
- Honda, S., Yuen, D. A., Balachandrar, S. & Reuteler, D. *Science* (in the press).
- Ito, E. & Takahashi, E. *J. geophys. Res.* **94**, 10637-10646 (1989).
- Shearer, P. M. & Masters, T. G. *Nature* **355**, 791-796 (1992).
- Dziewonski, A. M. & Anderson, D. L. *Phys. Earth planet. Inter.* **25**, 297-356 (1981).
- Osako, M. & Ito, E. *Geophys. Res. Lett.* **18**, 239-242 (1991).
- Anderson, D. L. *Phys. Earth planet. Inter.* **45**, 307-323 (1987).
- Ito, E., Akaogi, M., Topor, L. & Navrotsky, A. *Science* **249**, 1275-1278 (1990).
- Anderson, O. L., Oda, H. & Isaak, D. *Geophys. Res. Lett.* **19**, 1987-1990 (1992).
- Schubert, G. *Ann. Rev. Earth Planet. Sci.* **7**, 289-342 (1979).
- Schubert, G., Stevenson, D. J. & Cassen, P. *J. geophys. Res.* **85**, 2531-2538 (1980).
- Davies, G. F. & Richards, M. A. *J. Geol.* **100**, 151-206 (1992).
- Sharpe, H. N. & Peltier, W. R. *Geophys. Res. Lett.* **5**, 737-740 (1978).
- Houseman, G. *Nature* **332**, 346-349 (1988).
- Travis, B., Weinstein, S. & Olson, P. *Geophys. Res. Lett.* **17**, 243-246 (1990).
- Christensen, U. & Yuen, D. *Geophys. Res. Lett.* **15**, 597-600 (1988).
- Busse, F. H. in *Mantle Convection: Plate Tectonics and Global Dynamics* (ed. Peltier, W. R.) 23-95 (Gordon and Breach, New York, 1989).

34. Turcotte, D. L. & Schubert, G. *Geodynamics: Applications of Continuum Physics to Geological Problems* 279–285 (Wiley, New York, 1982).
35. Inoue, H., Fukao, Y., Tanabe, K. & Ogata, Y. *Phys. Earth planet. Inter.* **59**, 294–328 (1990).
36. Creager, K. C. & Jordan, T. H. *J. geophys. Res.* **91**, 3573–3589 (1986).
37. Vidale, J. E. & Garcia-Gonzales, D. *Geophys. Res. Lett.* **15**, 369–372 (1988).
38. Zhou, H. W. & Anderson, D. L. *Proc. natn. Acad. Sci. U.S.A.* **86**, 8602–8606 (1989).
39. Zhou, H. W. & Clayton, R. W. *J. geophys. Res.* **95**, 6829–6851 (1990).
40. van der Hilst, R., Engdahl, R., Spakman, W. & Nolet, G. *Nature* **353**, 37–43 (1991).
41. Fukao, T., Obayashi, M., Inoue, H. & Nenba, M. *J. geophys. Res.* **97**, 4809–4822 (1992).
42. Christensen, U. R. & Yuen, D. A. *J. geophys. Res.* **89**, 4389–4402 (1984).
43. Solheim, L. P. & Peltier, W. R. *J. geophys. Res.* (in the press).
44. Scrivner, C. & Anderson, D. L. *Geophys. Res. Lett.* **19**, 1053–1056 (1992).

45. Davies, G. F. *J. geophys. Res.* **93**, 10451–10466 (1988).
46. Gurnis, M. & Zhong, S. *Geophys. Res. Lett.* **18**, 581–584 (1991).
47. Gurnis, M. *Nature* **332**, 695–699 (1988).
48. Weinstein, S. A. *Geophys. Res. Lett.* **20**, 101–104 (1993).
49. Jordan, T. H., Puster, P., Glatzmaier, G. A. & Tackley, P. J. *Nature* (submitted).

ACKNOWLEDGEMENTS. We thank W. R. Peltier, D. L. Anderson, D. A. Yuen, M. Gurnis, V. Solomatov and the participants of the Los Alamos Mantle Convection Workshop for discussions, U. Christensen for his review and H. X. Qian for photographic assistance. The Intel Touchstone Delta System was operated by the California Institute of Technology on behalf of the Concurrent Supercomputing Consortium. Access to this facility was provided by the institute. P.T. thanks the NSF for financial support.

LETTERS TO NATURE

The proper motion of Geminga's optical counterpart

G. F. Bignami*†, P. A. Caraveo* & S. Mereghetti*

* Istituto di Fisica Cosmica del CNR, Via Bassini 15,
20133 Milano, Italy

† Dipartimento di Ingegneria Industriale, Università di Cassino,
Via Zamosh 43, 03043 Cassino, Italy

THE strong γ -ray source Geminga was first observed by the satellite SAS-2^{1,2} and later seen by the COS-B satellite^{3,4}. An association with the peculiar X-ray source 1E0630+178 was proposed⁵, suggesting that Geminga is a nearby neutron star (~ 100 pc from Earth) which is not visible as a radio pulsar. Searches for its optical counterpart yielded as the best candidate a very faint ($m_v = 25.5$) object, G'', which was proposed on the basis of its colours^{6–8}. The association of Geminga with 1E0630+178 was confirmed recently by the discovery of a 237-ms periodicity in the soft X-ray emission⁹ from the latter; such oscillations were recognized immediately afterwards in data¹⁰ from the Gamma Ray Observatory, and in reanalysis of the COS-B¹¹ and SAS-2 data¹². As its timing and energy parameters indicated that Geminga is closer than 400 pc, Bignami and Caraveo suggested¹¹ that the proper motion of the optical counterpart G'' might be measurable. Here we compare a 1992 optical image of G'' with previous data from 1984 and 1987, and find that the proper motion is 0.17 arcsec per year. This motion is consistent with the identification of G'' as a neutron star at a distance of about 100 pc, thus confirming it as the optical counterpart of Geminga.

On 4 November 1992 an observing run in service mode was made with the New Technology Telescope (NTT) of the European Southern Observatory (ESO) on La Silla, Chile. Using the Super Seeing Imager (SUSI)¹³ set up to take advantage of the very good seeing (0.6–0.8''), 10 exposures of 15' each were taken through a Johnson V filter, with a pixel size corresponding to 0.13''. The data were reduced and summed directly on the mountain (A. Moneti, personal communication). The resulting image was compared with two others of the same field, one obtained at the Canada-France-Hawaii Telescope (CFHT) 3.5-m instrument in January 1984 (ref. 6) and the other at the ESO La Silla 3.6-m in January 1987 (ref. 8).

Figure 1 shows a composite of the three images, where the 1984 and 1987 images have been rebinned and tilted to match the scale and orientation of the NTT 1992 frame. The motion of G'' to the northeast is apparent, showing a displacement of about 12 pixels between 1984 and 1992, with the 1987 data at the correct angle and position.

To assess the effect more quantitatively, and to exclude any possible systematic error, we adopted the following procedure. The pixel positions of 16 objects of magnitude ~ 24 –26 were measured in each of the three frames shown (partially) in Fig. 1. For each star, a mean position was calculated, and deviations from the mean are given in Fig. 2 in units of pixels (0.13'') for the x and y axes. Whereas the comparison objects seem to lie

well within the centring errors in each image, G'' does not. Its total displacement of 12 pixels to the northeast, which consists of two similar components ($\Delta x \approx 10$ px, $\Delta y \approx 7$ px), indicates the high confidence level of the observed effect, especially when compared with the average <2 pixel position jitter of the comparison faint stars. Furthermore, the three positions found for G'' show a coherent pattern of movement.

The coordinates of G'' at the three epochs are: for 1992, $\alpha_{(1950)} = 6$ h 30 min 59.15 s, $\delta_{(1950)} = 17^\circ 48' 33.6'' \pm 0.16''$; for 1987, $\alpha_{(1950)} = 6$ h 30 min 59.10 s, $\delta_{(1950)} = 17^\circ 48' 33.0'' \pm 0.68''$; for 1984, $\alpha_{(1950)} = 6$ h 30 min 59.06 s, $\delta_{(1950)} = 17^\circ 48' 32.7'' \pm 0.46''$. All the above coordinates have been computed in the original (not rebinned) data, using a common set of reference stars extracted from the original Hubble Space Telescope (HST) Guide Star Catalogue. The quoted uncertainties take into account the r.m.s. of the astrometry fit (0.10'', 0.12'' and 0.19'' in the 1992, 1987 and 1984 data, respectively) and the error in the centring of G'' (~ 1 pixel of each original data set: 0.13'', 0.675'', 0.412'' in the 1992, 1987 and 1984 images, respectively). These uncertainties only reflect the relative astrometry errors. As such, they cannot be used directly for absolute positions, because they do not take into account the absolute astrometry errors of the HST Guide Star Catalogue. A linear fit to the above coordinates gives the following components for the proper motion of G'': $\mu_\alpha = 0.14'' \text{ yr}^{-1}$, $\mu_\delta = 0.10'' \text{ yr}^{-1}$ for a total of $\mu = 0.17'' \text{ yr}^{-1}$, to which a conservative error of $\pm 0.05'' \text{ yr}^{-1}$ should be attached.

The large proper motion of G'' ($m_v = 25.5$) can only be interpreted in two ways: either the object is a Solar System body (for example an asteroid or comet), or it is a subluminal, truly faint star. The first possibility cannot be discarded lightly, in view of the low ecliptic latitude of Geminga. Two strong arguments against it, however, are the extremely slow motion (for a Solar System body) at a large angle with the ecliptic plane, and the low probability of the event, in view of the small area considered. Proof, in the end, will come from observing the proper motion of the X-ray source, the reference for which is the Einstein Observatory's High Resolution Imager (HRI) 1981 position. Comparison with the Rosat satellite's HRI 1991 position might be interesting but unlikely to be conclusive. Interpreting G'' as a star, the convenient relation $\mu = 0.2 v_{100} d_{100}^{-1}$ applies, where μ is in arcsec per year, v_{100} is the object velocity in the plane of the sky in units of 100 km s^{-1} , and d_{100} its distance in units of 100 pc. For a displacement similar to that observed, this gives a distance of the order of 100 pc for a velocity in the plane of the sky of 100 km s^{-1} , not far from the mean for the broad velocity distribution of radio pulsars¹⁴. At 100 pc, for example, the object would have a visible magnitude $M_v = 20.5$. This should be compared, for example, with $M_v = 15$ for the Vela pulsar; that is, it has an underluminosity only comprehensible for a neutron star. Anything more luminous would have to be correspondingly more distant and thus faster. For comparison, the Vela pulsar, at an accepted distance of 450 pc, has been measured¹⁵ to have a proper motion of about $0.05'' \text{ yr}^{-1}$. The only normal stars that could in principle mimic the magnitude and velocity of G'' are the 'late-type extreme subdwarfs', characterized by large transverse velocities, red colours and absolute

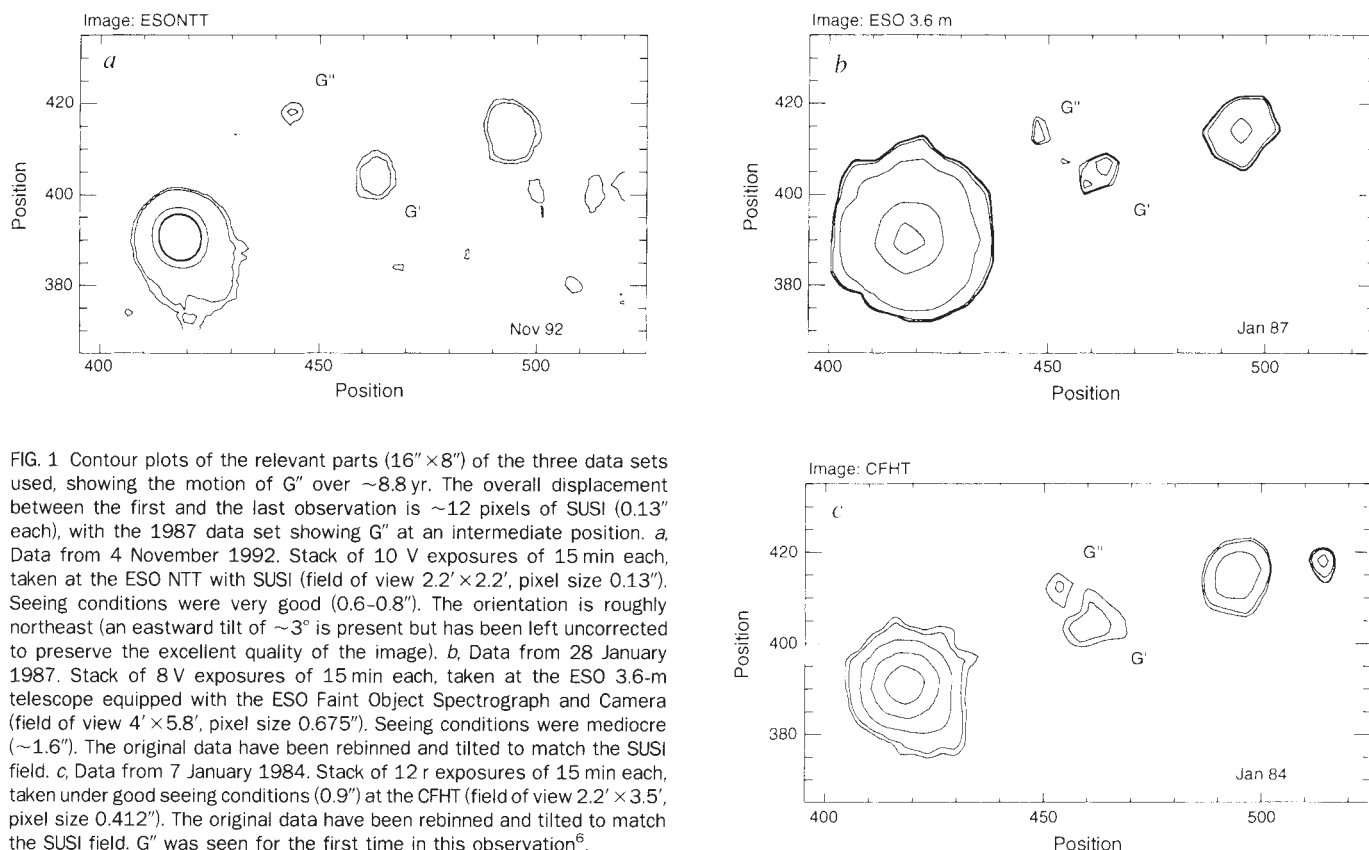


FIG. 1 Contour plots of the relevant parts ($16'' \times 8''$) of the three data sets used, showing the motion of G'' over ~ 8.8 yr. The overall displacement between the first and the last observation is ~ 12 pixels of SUSI ($0.13''$ each), with the 1987 data set showing G'' at an intermediate position. *a*, Data from 4 November 1992. Stack of 10 V exposures of 15 min each, taken at the ESO NTT with SUSI (field of view $2.2' \times 2.2'$, pixel size $0.13''$). Seeing conditions were very good (0.6 – $0.8''$). The orientation is roughly northeast (an eastward tilt of $\sim 3^\circ$ is present but has been left uncorrected to preserve the excellent quality of the image). *b*, Data from 28 January 1987. Stack of 8 V exposures of 15 min each, taken at the ESO 3.6-m telescope equipped with the ESO Faint Object Spectrograph and Camera (field of view $4' \times 5.8'$, pixel size $0.675''$). Seeing conditions were mediocre ($\sim 1.6''$). The original data have been rebinned and tilted to match the SUSI field. *c*, Data from 7 January 1984. Stack of 12 r exposures of 15 min each, taken under good seeing conditions ($0.9''$) at the CFHT (field of view $2.2' \times 3.5'$, pixel size $0.412''$). The original data have been rebinned and tilted to match the SUSI field. G'' was seen for the first time in this observation⁶.

magnitudes as faint as $M_v = 15$ (ref. 16). The apparent magnitude of G'' would place it at ~ 1 kpc, implying a velocity in excess of $1,000 \text{ km s}^{-1}$, a value far above any observed so far. This possibility, unlikely in view of the very small area considered, is ruled out by the bluer colour of G'' . Altogether, no known object other than a neutron star can explain the properties of G'' . The observed motion thus confirms the identification of G'' as the optical counterpart of Geminga, which has been shown to be a neutron star by the X-ray and γ -ray periodicity.

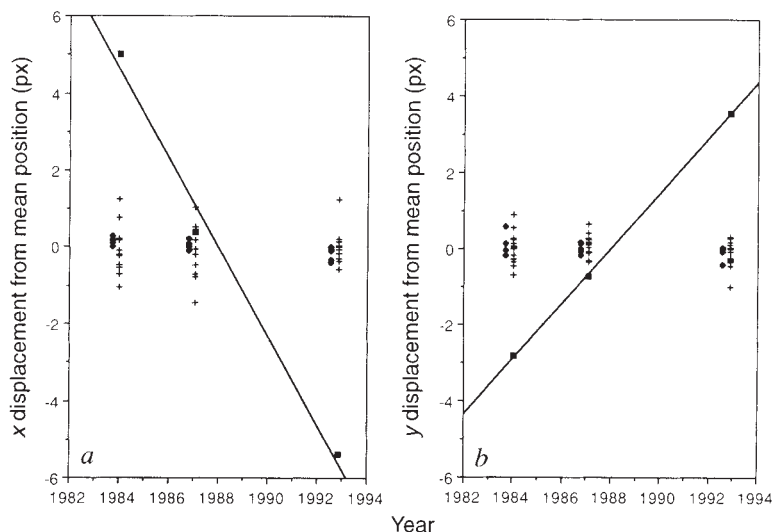
Geminga becomes the third neutron star identified in the optical, after the Crab and Vela pulsars. The pulsar 0540–69 in the Large Magellanic Cloud, although definitely seen to

pulsate at optical wavelengths¹⁷, has so far only a probable identification through imaging¹⁸.

The observed motion of G'' could help to explain difficulties encountered with the timing parameters of the object¹⁹. In particular, the second derivative of the period, when computed over a long time history to include both GRO and COS-B data (1975–1991), might be affected by changes in position as well as by period glitches.

It is interesting to speculate on the birthplace of Geminga, now that its direction of motion and angular velocity are known. If its age is really 340,000 years, as suggested by its period derivative¹¹, the object comes from a point about 16° roughly

FIG. 2 x and y displacements (in units of $0.13''$ pixels) from the mean position of G'' ($M_v = 25.5$, filled squares) together with 16 faint comparison objects, the magnitudes of which range from $M_v = 23.5$ to $M_v = 26.5$. These are stars G' , 31, 130, 132, 133, 137, 138, 139, 140, 146 and 155 (following the numbering given in ref. 7), plus other fainter ones visible in our three data sets. For each object, the positions obtained in the three frames were averaged and the three x and y deviations from the mean were computed (*a* and *b*, respectively). Diamonds represent five objects between $M_v \approx 23.5$ and $M_v \approx 25$ and crosses represent 11 objects fainter than $M_v \approx 25$. For clarity, diamonds are plotted slightly to the left of crosses, but measurements are contemporary. No systematic motion is present for the comparison objects, whose deviations appear to lie well within the centring errors (typically < 2 pixels) in each image. In contrast, G'' (filled squares) shows a clear northeast displacement. The linear fits to the x and y displacements of G'' from the mean are also shown.



to the southwest of its present position. This is well outside the boundaries of the Gemini constellation, so that the present name could not have been given at birth. In view of the small distance involved, it is probably difficult to search for a supernova remnant which may well now include the Earth. But if Geminga was generated in a supernova explosion liberating a huge amount of energy at 100 pc or so from the Earth, the environmental effects could have been considerable.

As to the physical nature of the optical emission, these data add no information, except for the possible constraint on Geminga's absolute magnitude. Thermal as well as nonthermal mechanisms may contribute to the emission (see, for example, ref. 20), which could be pulsed at 237 ms, with an unknown duty cycle. The period-luminosity dependence originally proposed by Pacini²¹, and discussed more recently in ref. 22, can be applied, assuming for Geminga an optical duty cycle similar

to that of the Vela pulsar, as is the case at higher energies. This yields an $M_v \approx 28$ which would place Geminga at ~ 3 pc, a distance corresponding to an improbable space velocity vector. From the γ -ray data and from the pulsar's measured rotational energy loss, an independent absolute upper limit of ~ 400 pc distance can be estimated.

An image of Geminga with the planetary camera on the HST is planned for the near future, to pinpoint the position of G' to the best possible accuracy allowed by its current point spread function. Repeating such measurements six months apart might lead to a parallax measurement for Geminga, something which cannot be easily done from the ground for such a faint object. The predictable direction of the annual parallactic displacement will help to disentangle it from the known proper motion, and its magnitude will solve the problem of a precise distance measurement. $\square$

Received 19 November 1992; accepted 29 January 1993.

1. Fichtel, C. E. *et al. Astrophys. J.* **198**, 163–182 (1975).
2. Thompson, D. J. *et al. Astrophys. J.* **213**, 252–262 (1977).
3. Swanenburg, B. N. *et al. Astrophys. J.* **243**, L69–L73 (1981).
4. Bignami, G. F. & Hermsen, W. *Ann. Rev. Astr. Astrophys.* **21**, 67–108 (1983).
5. Bignami, G. F., Caraveo, P. A. & Lamb, R. C. *Astrophys. J.* **272**, L9–L13 (1983).
6. Bignami, G. F., Caraveo, P. A., Paul, J. A., Salotti, L. & Vigroux, L. *Astrophys. J.* **319**, 358–361 (1987).
7. Halpern, J. H. & Tytler, D. *Astrophys. J.* **330**, 201–217 (1988).
8. Bignami, G. F., Caraveo, P. A. & Paul, J. A. *Astr. Astrophys.* **202**, L1–L4 (1988).
9. Halpern, J. P. & Holt, S. S. *Nature* **357**, 222–224 (1992).
10. Bertsch, D. L. *et al. Nature* **357**, 306–307 (1992).
11. Bignami, G. F. & Caraveo, P. A. *Nature* **357**, 287 (1992).
12. Mattox, J. R. *et al. Astrophys. J.* **401**, L23–L26 (1992).

13. Dekker, H., D'Odorico, S. & Melnick, J. *ESO Operating Manual*, 14 (1991).
14. Lyne, A. G., Anderson, B. & Salter, M. J. *Mon. Not. R. astr. Soc.* **201**, 503–513 (1982).
15. Ogelman, H., Koch-Miramond, L. & Auriere, M. *Astrophys. J.* **342**, L83–L86 (1989).
16. Monet, D. G. *et al. Astrophys. J.* **103**, 638–652 (1992).
17. Middleditch, J., Pennypacker, C. R. & Burns, M. S. *Astrophys. J.* **315**, 142–148 (1987).
18. Caraveo, P. A., Bignami, G. F., Mereghetti, S. & Mombelli, M. *Astrophys. J.* **395**, L103–L105 (1992).
19. Mayer-Hasselwander, H. A. *et al. IAU Circ. No.* 5649 (1992).
20. Curtis-Michel, F. in *Theory of Neutron Stars Magnetosphere* (Univ. of Chicago Press, 1991).
21. Pacini, F. *Astrophys. J.* **163**, L17–L19 (1971).
22. Pacini, F. & Salvati, M. *Astrophys. J.* **321**, 447–449 (1987).

ACKNOWLEDGEMENTS. We thank A. Smette for the NTT observations on 4 November, and D. Golombek for supplying the HST Guide Star Catalogue.

The Geminga supernova as a possible cause of the local interstellar bubble

Neil Gehrels & Wan Chen

Laboratory for High Energy Astrophysics,
NASA Goddard Space Flight Center, Greenbelt, Maryland 20771, USA

THE Solar System resides at the edge of a cavity of hot (10^6 K), low-density ($5 \times 10^{-3} \text{ cm}^{-3}$), X-ray emitting gas embedded in the interstellar medium^{1–4}. This void, sometimes called the Local Bubble, is thought to be less than 10^7 years old, but its origin is unknown. Here we propose that the void was caused by the supernova that produced the Geminga pulsar. The initial identification⁵ of Geminga as a pulsar, and the subsequent detection^{6–8} of pulsations in high-energy γ -rays, give an age of 3×10^5 years and a pulsar distance in the range 40 to 400 pc (refs 6, 7). Using this information, and the recently discovered^{9,10} proper motion of a likely optical counterpart, we find that the supernova was well positioned to produce the local void, provided that the explosion occurred within about 60 pc of the Solar System. Larger distances are not excluded by our analysis, but they would put the supernova at a position for which there is no evidence for such an energy input.

Geminga, discovered by SAS-2 in the early 1970s¹¹ and subsequently detected by COS-B¹², is one of the brightest high-energy (>100 MeV) γ -ray sources in the sky, but its nature remained a mystery for 20 years. During this period a tentative X-ray counterpart, 1E0630+178, was found¹³, as well as three possible faint optical counterparts denoted G (refs 14, 15), G' (refs 16, 17) and G'' (refs 18, 19). The Geminga mystery was solved in 1992 when 1E0630+178 was from Rosat X-ray data found to be a pulsar. The 237-ms period was then detected in the γ -ray emission from Geminga using data from GRO/EGRET⁶, COS-B⁷ and SAS-2⁸.

Pulsars are neutron stars born at the core of type II supernovae. The $\sim 1 M_\odot$ (1 solar mass) iron core of a massive progenitor star collapses to form a rapidly spinning and highly

magnetized neutron star, while the $>8M_\odot$ outer envelope is ejected with high speed ($>1,000 \text{ km s}^{-1}$). In some cases, slight asymmetries in the explosion cause the neutron stars to be ejected at a speed of $>100 \text{ km s}^{-1}$ (ref. 20).

The Geminga pulsar turns out to be young and remarkably nearby. The characteristic age, calculated by dividing the period by twice the period derivative, is 3.2×10^5 years based on the EGRET ephemeris⁶ and 3.4×10^5 years from the COS-B ephemeris^{8,21}. The distance is harder to determine, but can be estimated by comparing the observed γ -ray intensity with the luminosity expected from the available spin-down energy. Uncertainties come in from an unknown beaming factor and unknown γ -ray efficiency, but reasonable values range from <40 to 400 pc (refs 6, 7).

The recent discovery^{9,10} of proper motion in the G' candidate optical counterpart of Geminga (the most likely of the three candidates¹⁸) allows us to trace back in time and estimate the location of the supernova and also gives a check on the distance. A proper motion $0.17'' \pm 0.04''$ per year was measured with components $0.10''$ north in declination and $0.14''$ east in right ascension (G. Bignami, personal communication). The observed motion corresponds to a pulsar transverse speed of 78 km s^{-1} at 100 pc and 314 km s^{-1} at 400 pc, so the distance to Geminga is comfortable at 100 pc and is probably not much greater than 400 pc (ref. 20). The fact that these distances fall in the range estimated from the γ -ray intensity supports the G' identification.

The current position for the Geminga pulsar is $\alpha(2000) = 6 \text{ h } 34 \text{ min}$, $\delta(2000) = 17^\circ 46'$ (refs 13, 18) ($l^{\text{II}} = 194.8^\circ$, $b^{\text{II}} = 3.7^\circ$). Propagating backwards 3.3×10^5 years in the opposite direction to the proper motion gives a position for the Geminga supernova of $\alpha = 5 \text{ h } 40 \text{ min}$, $\delta = 8^\circ 24'$ ($l^{\text{II}} 197^\circ$, $b^{\text{II}} = -11.7^\circ$). The typical uncertainty radius is 3° due to uncertainties in the proper motion and age. Figure 1 shows the possible positions for the Geminga supernova (striped region) in radial projection on the galactic plane along with contours of hydrogen column density in the interstellar medium (ISM) from ref. 3. The direction is as given above ($l^{\text{II}} = 194.8^\circ$) with a 6° spread. The radial distance is constrained only to be greater than ~ 10 pc, as a closer supernova would have severely disrupted the Earth's atmosphere²², and less than several hundred parsecs given the γ -ray brightness of

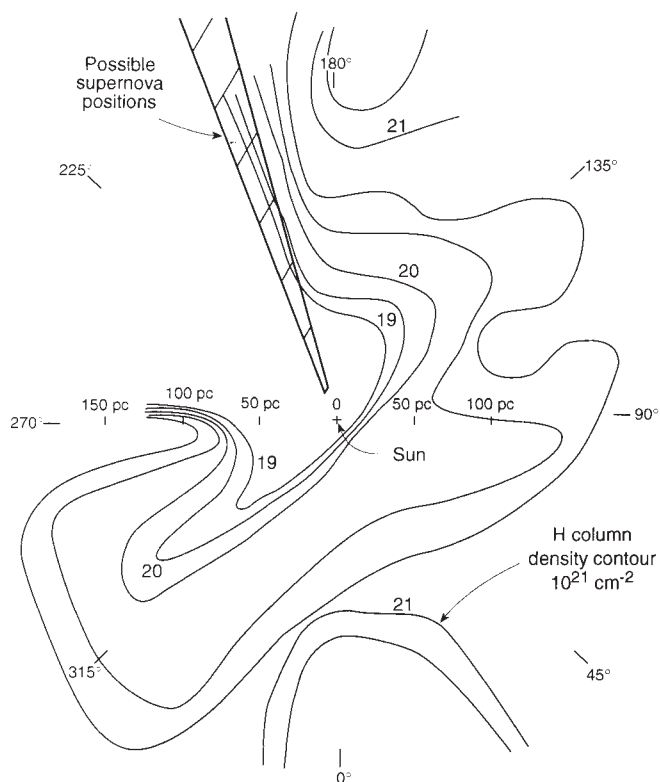


FIG. 1 Possible sites of the Geminga supernova 3.3×10^5 yr ago, shown in the striped region, projected on the galactic plane. The angular coordinates are in galactic longitude, where the Galactic Centre is towards 0° . Superimposed are the H I column density contours surrounding the Local Bubble from Paresce³ ($N(\text{H I}) = 10^{19}, 5 \times 10^{19}, 10^{20}, 2 \times 10^{20}, 5 \times 10^{20}, 10^{21}, 2 \times 10^{21} \text{ cm}^{-2}$). The Local Bubble occupies most of the open space in the upper left region between the high density walls and forms an elongated structure perpendicular to the galactic plane. The Per association is towards 145° and behind the $5 \times 10^{20} \text{ cm}^{-2}$ contour. The Sco-Cen associations are towards 320° behind the 10^{20} cm^{-2} contour.

the pulsar. Typical pulsar radial velocities of $100\text{--}300 \text{ km s}^{-1}$ over 3×10^5 years imply that the supernova distance may be $50\text{--}100 \text{ pc}$ different from that of the pulsar. The supernova direction in the figure is close to the current direction to Geminga because most of the proper motion is in galactic latitude (pulsar $b^{\text{II}} = 3.7^\circ$, supernova $b^{\text{II}} = -11.7^\circ$).

The local cavity is apparent in Fig. 1. It has an opening towards 225° and a large-scale high-density boundary with the Perseus OB associations at 145° and the Scorpius-Centaurus association at 320° . Perpendicular to the galactic plane, the cavity extends $50\text{--}100 \text{ pc}$ south and $>100 \text{ pc}$ north, forming an elongated galactic 'chimney'. The Solar System is in fact not

located directly in the Local Bubble, but is embedded in a warm ($10^3\text{--}10^4 \text{ K}$) and medium-density (0.1 cm^{-3}) cloud of $\sim 20 \text{ pc}$ in size and moving away slowly ($\sim 20 \text{ km s}^{-1}$) from the nearby wall. It has long been a puzzle how this bubble and warm fringe were formed in the local ISM³. We propose that the Geminga supernova is a natural explanation for all the observed structure.

It is clear from Fig. 1 that the Geminga supernova was well positioned to create the Local Bubble only if its radial distance from the Solar System was less than $\sim 60 \text{ pc}$. The lack of observed disturbances of the ISM anywhere else along the band of possible supernova sites is strong evidence that the supernova was nearby. It has been pointed out that a single supernova could be energetically adequate by itself to create the Local Bubble and fill it with hot gas if the initial ISM density in the region was relatively low ($\sim 10^{-2} \text{ cm}^{-3}$, ref. 1). In that case, the energy to produce the bubble is $\approx 3 \times 10^{50} \text{ erg}$, which, added to the thermal energy of the 10^6 K gas ($\sim 3 \times 10^{50} \text{ erg}$), is within the range of the kinetic energy of a typical type II supernova, $\sim 10^{51} \text{ erg}$. Although it may have been hard for the ejecta to propagate against the dense shell in the direction of the Sun, it was relatively easy for it to push upwards from the disk, where the density drops exponentially, to form the observed galactic 'chimney'. In the galactic plane, the Solar System is located at the inner edge of the Orion spur which stretches between the Sagittari-Carina arm and the Perseus arm²³. Although the direction of 225° points towards a low density region of space, the density structure in the local ISM is probably more relevant to the formation of the opening in the Local Bubble.

After the explosion, the interaction of the expanding remnant with other remnants and initial mass distributions in the ISM evidently caused the pocket and steep contours near the Sun. The expansion stopped at the nearby cold, dense wall and all the kinetic energy was converted to thermal energy. The outer layer of the wall was evaporated by gas at 10^6 K (ref. 24) and formed the observed warm cloud surrounding the Solar System. The warm gas expansion velocity of $\sim 20 \text{ km s}^{-1}$, and the total mass evaporated during the last $3.3 \times 10^5 \text{ yr}$, $\sim 100 M_\odot$ according to the evaporation theory²⁴, are fully consistent with the observed values³. The warm and hot regions have now reached pressure equilibrium.

The Geminga supernova is a unique phenomenon in astrophysics. As a type II supernova occurring 3×10^5 years ago at a distance of 60 pc it would have been as bright ($m \approx -13$) as the Moon. It created (1) a pulsar that is one of the brightest sources in the γ -ray sky, (2) a void filled with hot gas in the local ISM that permits extreme ultraviolet observations of nearby stars, (3) a galactic chimney above the disk and an opening in the local ISM and (4) a three-component interstellar medium structure in the solar neighborhood.

Note added in proof: New studies of the X-ray spectrum and flux indicate that the Geminga pulsar distance is between 150 and 400 pc ²⁵, which implies a radial velocity of $>250 \text{ km s}^{-1}$ for a supernova distance of $<60 \text{ pc}$. $\square$

Received 25 November 1992; accepted 11 January 1993.

1. Cox, D. P. & Reynolds, R. J. *A. Rev. Astr. Astrophys.* **25**, 303–344 (1987).
2. Bochkarev, N. G. *Astrophys. Space Sci.* **138**, 229–302 (1987).
3. Paresce, F. *Astr. J.* **89**, 1022–1037 (1984).
4. McCammon, D. & Sanders, W. T. *A. Rev. Astr. Astrophys.* **28**, 657–688 (1990).
5. Halpern, J. P. & Holt, S. S. *Nature* **357**, 222–224 (1992).
6. Bertsch, D. L. *et al. Nature* **357**, 306–307 (1992).
7. Bignami, G. F. & Caraveo, P. A. *Nature* **357**, 287–287 (1992).
8. Mattox, J. R. *et al. Astrophys. J.* (in the press).
9. Bignami, G. F., Caraveo, P. A. & Mereghetti, S. *IAU Circ. No.* 5651 (1992).
10. Bignami, G. F., Caraveo, P. A. & Mereghetti, S. *Nature* **361**, 704–706 (1993).
11. Fichtel, C. E. *et al. Astrophys. J.* **198**, 163–182 (1975).
12. Bennett, K. *et al. Astr. Astrophys.* **56**, 469–471 (1977).
13. Bignami, G. F., Caraveo, P. A. & Lamb, R. C. *Astrophys. J.* **272**, L9–L13 (1983).
14. Caraveo, P. A., Bignami, G. F., Vigroux, L. & Paul, J. A. *Astrophys. J.* **276**, L45–L47 (1984).
15. Sol, H., Tarengi, M., Vanderriest, C., Vigroux, L. & Lelievre, G. *Astr. Astrophys.* **144**, 109–114 (1985).

16. Djorgovski, S. & Kulkarni, S. *Astr. J.* **91**, 90–97 (1986).
17. Bignami, G. F., Caraveo, P. A., Paul, J. A., Salotti, L. & Vigroux, L. *Astrophys. J.* **319**, 358–361 (1987).
18. Halpern, J. P. & Tytler, D. *Astrophys. J.* **330**, 201–217 (1988).
19. Bignami, G. F., Caraveo, P. A. & Paul, J. A. *Astr. Astrophys.* **202**, L1–L4 (1988).
20. Harrison, P. A., Lyne, A. G. & Anderson, B. in *X-ray Binaries and Recycled Pulsars* (eds van den Heuvel, E. P. J. & Rappaport, S. A.) 155–160 (Kluwer, Dordrecht, 1992).
21. Hermsen, W. *et al. IAU Circ. No.* 5541 (1992).
22. Ruderman, M. *Science* **184**, 1079–1081 (1974).
23. Gilmore, G., King, I. R. & van der Kruit, P. C. *The Milky Way as a Galaxy* (Univ. Science Books, Mill Valley, 1990).
24. McKee, C. F. & Ostriker, J. P. *Astrophys. J.* **218**, 148–169 (1977).
25. Halpern, J. P. & Ruderman, M. *Astrophys. J.* (submitted).

ACKNOWLEDGEMENTS. We thank G. Bignami for providing us with Geminga proper motion components and for discussions. We appreciate discussions with F. Bruhweiler, E. Dwek, J. Halpern, F. Paresce, D. Thompson and J. Tueller.

Locating global minima in optimization problems by a random-cost approach

Bernd A. Berg*

Institute for Advanced Study Berlin, Wallotstrasse 19, 1000 Berlin 33, Germany

OPTIMIZATION problems are common to many diverse disciplines. The classic example is the travelling-salesman problem^{1,2}, in which the objective is to find the shortest path connecting a number of cities. Spin glasses³, meanwhile, exemplify a general class of systems subject to conflicting constraints: in this case, the impossibility of each spin in a lattice aligning favourably with all of its neighbours leads to 'frustration' and to a large number of local energy minima. The optimization problem then becomes the attempt to find the global minimum, or ground state. Predicting the conformational ground state of proteins⁴ is closely allied to the spin-glass problem. The chief difficulty in searching for global minima is that straightforward search algorithms¹ tend to become trapped in local minima. Only a few promising approaches, such as simulated annealing² or genetic algorithms⁵, exist. Here I present a general purpose Monte Carlo procedure in which local redirections of the search path are effected at all relevant length scales while enforcing a one-dimensional random walk in the function being minimized. This method yields a series of extrema, from which the global minimum can be extracted with high probability in the limit of large statistics.

In the following I refer to the function being minimized, $f = f(x_1, \dots, x_d)$, as the cost function. In the travelling-salesman problem, this is related to the economic cost of the journey. The simulated-annealing approach identifies the cost function with the energy of a statistical mechanical system at some temperature T , and tries to avoid trapping of the search in local minima by slow cooling accompanied by 'thermal' fluctuations. It is widely used with considerable success, but nevertheless suffers from a number of shortcomings. In particular, the method requires subtle optimization of the annealing schedule (defined by parameters such as the number of re-starts, the size of the temperature steps during annealing, and the temperature range). The algorithm that I introduce here promises to circumvent these shortcomings while maintaining or even improving the efficiency.

As in simulated annealing, a procedure must be defined for effecting local rearrangements or redirections of the search subsequently called updates. Simulated annealing accepts or rejects proposed updates with probabilities determined by the Metropolis criteria⁶. In contrast, in my approach updates are accepted or rejected in such a way that a random walk in the cost function $f(x_1, \dots, x_d)$ becomes enforced (this should not be confused with a random path in the arguments $x_1, \dots, x_d$). The idea is that the random walk in f will avoid trapping in a local minimum and, less obviously, will still provide a valid method to search for minima. To see this, let us introduce a bound, $f^{\max}$, and constrain our interest in the cost function to values $f \leq f^{\max}$. We assume that the random path has a finite average step size $\bar{\Delta}f$, and define $M = (f^{\max} - f^{\min})/\bar{\Delta}f$, where the numerator is the typical difference between a local maximum ($f^{\max} \leq f^{\max}$) and a local minimum ($f^{\min} \geq f^{\min}$). It is well known that, with an appropriate constant c , the random walk will need an average of cM^2 steps to move from a maximum to a minimum or vice versa.

A similar random walk picture⁷ led to the introduction of the multicanonical ensemble⁸, for which the sampled energy density is flat (compare with ref. 9 for the earlier 'umbrella' sampling).

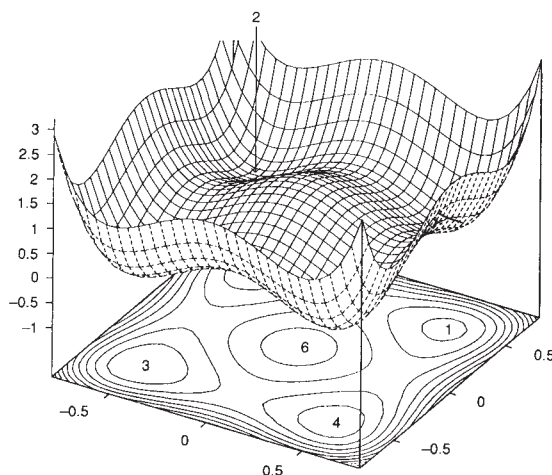


FIG. 1 Shape of the trial function. Equipotential lines are given in the underlying x_1 - x_2 plane. The minima are labelled 1, 2, 3 and 4, and the maximum is labelled 6. Here function arguments are restricted to $-0.8 \leq x_i \leq 0.8$, to avoid obscuring the relevant region.

These multicanonical simulations were successful in coping with first-order phase transitions⁸. A subsequent application^{10,11} to spin glasses made a connection with simulated annealing. Such simulations allowed efficient ground-state calculations. But the determination of the essential weight factors relied on a cumbersome recursive procedure¹⁰, maintaining an exact relationship with canonical statistical mechanics. Here I introduce a simple and general procedure to enforce the random walk.

Assume that the configuration $x_1, \dots, x_d$ is given, and let us generate update proposals

$$x_1, \dots, x_d \rightarrow x_1^n, \dots, x_d^n \quad (n=1, \dots, N) \quad (1)$$

A cost difference is associated with each proposal

$$\Delta_n f = f(x_1^n, \dots, x_d^n) - f(x_1, \dots, x_d) \quad (2)$$

The set of proposed updates, specified by $N = N(x_1, \dots, x_d)$ has to be chosen such that a sufficient overview of configurations which can be reached by the local update procedure is obtained. Next we divide our set into subsets of positive, zero and negative cost differences

$$\Delta_j^+ f > 0 \quad (j=1, \dots, J^+), \quad \Delta_j^0 f = 0 \quad (j=1, \dots, J^0) \quad (3a)$$

and

$$\Delta_j^- f < 0 \quad (j=1, \dots, J^-), \quad (3b)$$

where $J^+ + J^0 + J^- = N$. To avoid technical complications, let $J^0 = 0$ for the moment. If $J^- = 0$ ($J^+ = 0$) we have reached a local minimum (maximum), which we record before proceeding with one of the updates. The more usual case will be $J^\pm \neq 0$. Then we pick, according to some statistical algorithm, one member of the $\Delta_j^+ f$ and one member of the $\Delta_j^- f$ group. For each group we also define an average update difference: $\bar{\Delta}^+ f > 0$ and $\bar{\Delta}^- f < 0$. The $\bar{\Delta}^\pm f$ can either be taken as expectation values with respect to the statistical selection procedure, or be identified with the values actually chosen. In each case a random walk in the cost function is now enforced if one chooses the probabilities $P^\pm$ ($P^- + P^+ = 1$) of finally accepting the update from the minus or plus groups, respectively, so that they obey

$$-P^- \bar{\Delta}^- f = P^+ \bar{\Delta}^+ f \quad (4)$$

This yields $P^- = \bar{\Delta}^+ f / (\bar{\Delta}^+ f - \bar{\Delta}^- f)$. Accepting an update often causes only few changes in the table of cost differences (equation (2)). The reason is that in typical nonlinear systems, variables

* On sabbatical from the Florida State University, Tallahassee, Florida 32306, USA.

interact only with their nearest neighbours. In the case that $\Delta_j^0 f$ updates play a role, some freedom arises in assigning to them a probability of P^0 . This may be exploited to optimize the performance.

Each updating step by the outlined procedure requires no information beyond the currently stored system. Recursive updates with an algorithm with this property define a Markov process. Consequently, we have established a Markov process which moves the system along a random walk in the cost function. About every cM^2 steps we encounter another extremum in the chain of configurations. Minima separated by a local maximum or an approach to the imposed $f^{\max}$ are to a high degree statistically independent. Let us generate altogether m of those independent minima and let the total number of local minima be K^- . Under the assumption that the global minimum attracts the Markov process with at least as high a probability as does the average local minimum, a bound on the probability of missing the global minimum is $P_{\text{miss}} < (1 - 1/K^-)^m$. I suggest that under typical circumstances the actual performance will be much better than indicated by this bound. The reason is that nearby local minima can be explored without moving far away from the $\bar{f}^{\min}$ region. Details will depend on particular applications and implementations; for instance, one cannot *a priori* exclude bottlenecks in the configuration space or unusual autocorrelations, although at present it is hard to imagine circumstances under which this would happen.

In the two-dimensional Edwards-Anderson-Ising spin-glass model, the random-cost algorithm finds ground-state energies more efficiently than multicanonical simulations¹⁰. Here $P^0 = P^- J^0 / J^-$ (for $J^- \neq 0$) was chosen. I have also applied the method to search for ground-state conformations of the protein met-enkephalin, for which extensive experience with simulated annealing¹² and other methods¹³ exists. I modified the computer program of ref. 12 by changing the updating procedure from Metropolis to random-cost. The immediate performance of the random-cost method was similar to that of ref. 12, which had included a carefully tuned annealing schedule. Details of both cases will be published elsewhere.

For the present purposes I will demonstrate the method for a much simpler (indeed, trivial) example. As trial function let us consider

$$f(x_1, \dots, x_d) = A \sum_{i=1}^d (x_i^2 - B)^2 + C \sum_{i=1}^d x_i \quad (5)$$

in the range $-1 \leq x_i \leq +1$. The restriction to this range imposes implicitly a bound $f^{\max}$. If $C = 0$, there is a maximum at $x_i = 0$, and the combinations of $x_{0,i} = \pm\sqrt{B}$ lead to $K^- = 2^d$ minima.

TABLE 1 Selected computer data

Update	Property	x_1	x_2	Function
0	Start	0.33333333	0.33333333	0.45246914
186	Minimum 1	0.49492293	0.49492293	0.09949490
2,443	Minimum 2	-0.50492694	0.49492293	-0.00050010
3,841	Minimum 3	-0.50492694	-0.50492694	-0.10049510
5,303	Minimum 2	-0.50492694	0.49492293	-0.00050010
8,978	Minimum 3	-0.50492694	-0.50492694	-0.10049510
10,748	Minimum 4	0.49492293	-0.50492694	-0.00050010
11,059	Minimum 1	0.49492293	0.49492293	0.09949490
13,023	Minimum 3	-0.50492694	-0.50492694	-0.10049510
13,911	Minimum 4	0.49492293	-0.50492694	-0.00050010
21,267	Minimum 3	-0.50492694	-0.50492694	-0.10049510
22,623	Minimum 1	0.49492293	0.49492293	0.09949490

For our illustration we choose $d = 2$, $A = 10$, $B = 0.25$ and $C = 0.1$. The small C term is to lift the degeneracy of the minima. Figure 1 depicts the function shape.

Let us denote by ε the precision to which we desire to determine the locations of the extrema. Here I choose $\varepsilon = 2^{-27} = 7.45 \times 10^{-9} < 10^{-8}$. The following $N = 108$ changes of single function arguments define our updates (equation (1)):

$$x_1^n = x_1 + 2^{-n}, \quad x_2^n = x_2, \quad n = 1, \dots, 27$$

$$x_1^n = x_1 - 2^{-n+27}, \quad x_2^n = x_2, \quad n = 28, \dots, 54$$

$$x_1^n = x_1, \quad x_2^n = x_2 + 2^{-n+54}, \quad n = 55, \dots, 81$$

$$x_1^n = x_1, \quad x_2^n = x_2 - 2^{-n+81}, \quad n = 82, \dots, 108$$

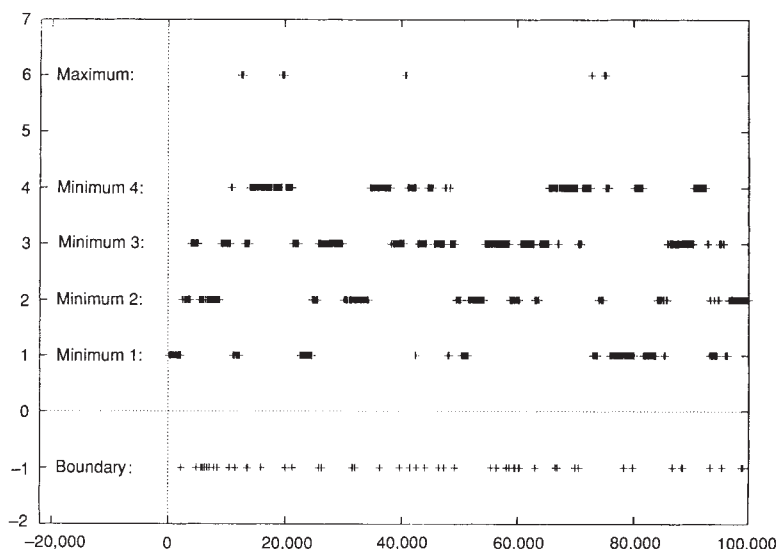
From this, the table of cost differences $\Delta_n f$ (2) is readily calculated. The choice of the update set is not unique. In this special situation, geometric progressions are a convenient means of achieving the required high accuracy. The subsets $\Delta_j^+ f$ ($j = 1, \dots, J^+$) and $\Delta_j^- f$ ($j = 1, \dots, J^-$) are determined next. (For this particular trial function the $\Delta_j^0 f$ set can be ignored.) As selection procedure, we adopt equal probabilities for updates within the same group. This gives rise to

$$\bar{\Delta}^\pm f = \frac{1}{J^\pm} \sum_{j=1}^{J^\pm} \Delta_j^\pm f$$

and the probabilities of choosing the $\Delta_j^- f$ rather than the $\Delta_j^+ f$ group follow from equation (4).

Starting with $x_1 = x_2 = \frac{1}{3}$, I have generated a Markov chain of 100,000 steps. Updates that would lead out of the range $-1 \leq x_i \leq 1$ are rejected. Figure 2 analyses the time series with respect to the encountered minima, the maximum and the boundary

FIG. 2 Random-cost time series with respect to boundary rejections, the four minima, and the maximum. Numbers are assigned to minima and maximum as in Fig. 1. Boundary rejections are indicated by -1. The finite extension of the marker causes some (non-existing) overlaps.



rejections. All extrema are visited. (The maximum is only visited a few times, but this is no surprise, as our trial function takes on larger values for a continuum of arguments on the boundary.) Table 1 collects computer data for the first 11 non-equal minima; the desired precision is indeed achieved.

The power of the random-walk constraint is best seen by comparison with naive sampling. With our accuracy $\varepsilon < 10^{-8}$ the range $-1 \leq x_i \leq +1$ is divided into a lattice of more than 4×10^{16} sites. The probability of hitting the global minimum at random is therefore $< 0.25 \times 10^{-16}$. On average, one encounters the global minimum about every 7,500 updates. Each update requires the recalculation of a table of 54 entries. Putting both together implies an improvement factor $> 10^{11}$. Of course, random choices of coordinates followed by conventional determinations of the local extrema would do better for our trivial trial function. The real strength of the random-cost method is that it still works well for frustrated systems. $\square$

Received 16 September 1992; accepted 5 January 1993.

1. Press, H. W., Flannery, B. P., Teukolsky, S. A. & Vetterling, W. T. *Numerical Recipes* (Cambridge Univ. Press, 1988).
2. Kirkpatrick, S., Gelatt, C. D. & Vecchi, M. P. *Science* **220**, 671–680 (1983).
3. Fisher, K. H. & Hertz, J. A. *Spin Glasses* (Cambridge Univ. Press, 1991).
4. Frauenfelder, H., Chu, K. & Philip, R. in *Biologically Inspired Physics*, NATO ASI Series **B263**, 1–14 (Plenum, New York, 1991).
5. Holland, J. H. *Scient. Am.* 44–50 (July 1992).
6. Metropolis, N., Rosenbluth, A., Rosenbluth, M., Teller, A. & Teller, E. *J. chem. Phys.* **21**, 1087–1092 (1953).
7. Berg, B. & Neuhaus, T. *Phys. Lett.* **B267**, 249–253 (1991).
8. Berg, B. *Int. J. mod. Phys. C3*, 1083–1096 (1992).
9. Torrie, G. M. & Valleau, J. P. *J. comput. Phys.* **23**, 187–199 (1977).
10. Berg, B. & Celik, T. *Phys. Rev. Lett.* **69**, 2292–2295 (1992).
11. Marinari, E. & Parisi, G. *Europhys. Lett.* **19**, 451–458 (1992).
12. Okamoto, Y., Kikuchi, T. & Kawai, H. *Chem. Lett.* **1992**, 1275–1278.
13. Nayem, A., Vila, J. & Scheraga, H. A. *J. comput. Chem.* **12**, 594–609 (1991).

ACKNOWLEDGEMENTS. I thank U. Hansmann for discussions and H. Kleinert for hospitality at the Freie Universität Berlin.

Climate-induced fluctuations in sea level during non-glacial times

David K. Jacobs* & Dork L. Sahagian†

* Department of Invertebrates, American Museum of Natural History, Central Park West at 79th Street, New York, New York 10024-5192, USA

† Department of Geological Sciences and Byrd Polar Research Center, Ohio State University, 155 South Oval Mall, Columbus, Ohio 43210-1397, USA

DURING periods of time when the Earth supports ice caps, sea level fluctuates periodically at intervals of 10^4 – 10^5 yr, with amplitudes of tens to more than 100 metres. These fluctuations result from expansions and contractions of continental ice sheets that occur in phase with the Milankovitch periodicities of the Earth's orbit^{1,2}. Smaller-amplitude sea-level fluctuations associated with Milankovitch periodicities are also evident during warm periods, such as the Late Triassic^{3–5}, which lack strong evidence for continental glaciation⁶. We argue here that, in times of limited ice volume, periodic climate-induced changes in lake and groundwater storage have the potential to produce small fluctuations in sea level. This mechanism could contribute a small component of the sea-level change observed in the Quaternary period, and may dominate the Milankovitch eustatic sea-level signal during earlier periods with limited ice volume⁶. We present evidence of contemporaneous Milankovitch periodicities in Late Triassic lake sediments⁷ and sea-level fluctuation^{3–5} which support the causal link between lake water storage and eustasy.

The mechanism for periodic storage of lake and groundwater is based on the Quaternary record of climate change. The strength and region of influence of the monsoon fluctuates with a periodicity of ~ 20 kyr as a consequence of precession of the

equinoxes^{8–10}. When the Earth is closer to the Sun (perigee) during the Northern Hemisphere summer, as it was $\sim 9,000$ years ago, global climate models indicate greater summertime heating and lower atmospheric pressure over the Asiatic land mass. These conditions result in greater monsoonal flow and increased summertime precipitation over North Africa and South Asia^{8–10}, an area encompassing more than 1/4 of the Earth's continental surface (Fig. 1).

To quantify the potential change in lake and groundwater storage, we determined the volumes of the 10 largest internally drained basins within the expanded region of influence of the early Holocene monsoon^{8–15}, and used these volumes to calculate a potential change in sea level (see Table 1 and Fig. 1). The Tarim basin alone, if filled with water, would lower sea level by 1 m. The volume of the 10 basins taken together is equivalent to 2.1 m of sea level.

Lakes represent the intersection of the water table with the ground surface, so lake level provides a proxy measure for groundwater storage. North and Central African lake basins¹¹ were at highstand $\sim 9,000$ years ago, as were lakes in South Asia^{10,12}. Before 10,000 years ago and after 5,000 years ago less than 20% of these lakes were at highstand¹¹. Although most of these basins are smaller than those in Table 1, many are associated with extensive aquifers such as the oasis complexes and the sand seas of the Sahara (Wadi Howar¹⁶, Selima Oasis¹⁷, Western Erg¹⁸, Oyo¹⁹, Grang Erg of Bilma²⁰). There are also many internally drained basins within the area of expanded monsoonal influence which have not been investigated in detail. These basins and their aquifers have large storage potential and the potential for rapid draw-down. Estimates of total groundwater on Earth vary, but they are of the same order of magnitude as continental ice volume²¹. Consequently, even small percentage changes in groundwater should produce considerable changes in sea level. Given the large fraction of the Earth's surface affected by orbitally forced changes in the monsoon, the smaller basins combined with their groundwater storage potential could have considerably more effect than the 10 basins considered (Table 1). A water volume of 2–6 m could be accommodated in smaller basins and aquifers in the region of monsoonal influence. Lake basins and groundwater combined could store monsoonal rains equivalent to ~ 4 to 8 m of sea level.

Quaternary and Holocene climates are complicated by the influence of glacial advance and retreat on global circulation^{22,23}. Consequently, the more northerly basins in Asia may not have filled completely or synchronously. The Caspian basin may have received glacial meltwater from the Volga and Turgai during Pleistocene/Holocene glacial retreat as well as receiving water from the Amu Darya and other drainages that should respond to expanded monsoonal precipitation²⁴. Drainages leading to Qinghai Lake and those leading to Tarim and Quidam are intimately associated. Qinghai Lake does show the influence of the expanded monsoon. Lake levels rose rapidly beginning 10,000 years ago and then declined after 6,000 years ago¹². Tarim and Quidam must have had a similar influx during this period, but it is not clear whether the Tarim basin filled completely. Our estimate of water storage is intended only as an indication of potential sea-level change in the past, not as an exact model of events in the Quaternary. This estimate also does not account for greater monsoonal intensity or larger storage potential associated with different continental configurations in the past.

Triassic palaeogeography seems to have been particularly amenable to monsoonal driven fluctuations in water storage. Triassic land masses were associated into one large continent (Pangea), much of which lay in subtropical to temperate latitudes. Sea level was generally low, resulting in large continental area, and a large seaway (Tethys) extended into the Pangean continent at tropical latitudes. These factors are thought to have enhanced monsoonal circulation, and widespread sedimentological evidence indicates a strongly seasonal climate²⁵. Global climate models based on Triassic continental configurations also

indicate strong monsoonal flow²⁶, leading to the term 'mega-monsoon' for this circulation pattern.

During the Triassic, incipient continental breakup produced a broad belt of graben and pull-apart basins extending from what is now the Gulf of Mexico along the east coast of North America to northern Europe and the Arctic^{27,28}. Lacustrine sediments fill nearly the entire Newark graben⁷, and lacustrine deposits, which often contain evidence of flooding and desiccation, are exposed from South Carolina to Nova Scotia²⁹ as well as in Morocco³⁰ and the Southwestern United States²⁵. Triassic evaporite deposits are the most extensive of any period, testifying to the frequent flooding and desiccation of many Triassic basins^{25,31}. Rift graben^{27,28}, lacustrine deposits^{7,25,29,30}, evaporite deposits^{25,31} and dune fields associated with lakes^{18-20,25} all had a greater areal extent in the Triassic than today, and these features were concentrated in the Northern Hemisphere subtropics, the region of precessionally forced fluctuations of monsoonal precipitation. Thus the Late Triassic may have shown more response to monsoonal fluctuation than could occur today.

Lake sediments of the Newark Supergroup form cyclic packages. Spectral analyses of these cycles, in the Carnian-age Lockatong formation and the overlying Norian-age Passaic formation⁷, indicate a dominant periodicity of ~20 kyr, the periodicity of equinox precession^{1,2} and changing monsoonal intensity⁸⁻¹⁵. Lower-frequency Milankovitch parameters were also observed, suggesting that orbital forcing of the monsoon resulted in fluctuating water flow into these basins⁷.

The Lofer sequences of the Dachstein in the northern Alps and the Latemar cycles from the Dolomite mountains of northern Italy are composed of nearshore marine carbonate rocks that exhibit metre-scale eustatic cycles³⁻⁵. The Norian Lofer cycles correlate with lacustrine beds of the Passaic formation. The Ladinian Latemar cycles immediately precede the Carnian Lockatong cycles⁵. The Lofer and Latemar sections consist of progradational packages with subarially exposed caps indicating repeated eustatic flooding of a continuously subsiding platform³⁻⁵. Spectral analyses³⁻⁵ suggest a dominant 20-kyr periodicity and bundling of these cycles into groups of five similar to the Milankovitch structure observed in the correlative lacustrine sediments⁷. These results are consistent with a causal

TABLE 1 The volume of empty basins

Basin	Area (10 ¹⁰ m ²)	Volume (10 ¹³ m ³)	Sea level (cm)	Outflow elevation
1 Tarim	87	39	108	1,500 m (northeast rim)
2 Caspian	232	13	36	100 m (Don River)
3 East Iran	57*	7*	19*	900 m (northwest)*
4 Afghan-Zabol				
5 Baluchistan				
6 Chad	67	5.9	16	400 m (Benue River)
7 Uvs Nor	22	5.2	14	1,600 m (Samaltagatay)
8 Quidam	16	3.3	9	3,500 m (Lapiquan)
9 Balkash-Zungar	30	1.9	5.3	400 m (Caspian)
10 Esfahan	5.1	5.8	1.6	1,800 m (northeast to east Iran)

* This figure is for east Iran, Afghan-Zabol and Baluchistan combined, which proved to be a complex interconnected basin.

Empty lake basin volumes were calculated using topographic data from the digital terrain file ET0P05U of the Defense Mapping Agency. Detailed topographic maps were constructed for each of the basins. Outflow elevations were determined from these and checked against published topographic maps. A Fortran program was used to sum over a set of area elements within the closed contour at the outflow elevation. Each area element was multiplied by a potential lake water depth determined by the difference between the topographic surface and the outflow elevation. These volumes, and values of equivalent sea-level change, do not include groundwater. Contributions to sea-level change of 2–6 m from additional basins and groundwater were modelled as a uniform layer of porous material through which the groundwater table passes up and down over the entire region of increased monsoonal influence. The actual porosities will vary widely, as will the depth through which the water table fluctuated. Within the region in question there are additional basins that could hold water; thick deposits of loess and sand with porosities often in excess of 0.4; Tertiary fluvial and lacustrine sediments and Mesozoic rocks with porosities in excess of 0.3; and smaller areas of basement rock with porosities of ~0.1. With a conservative average porosity of 0.3, an average rise in water table of 200 m would lower sea level by 6 m. A porosity of 0.3 in combination with an average increase in the water table of 65 m would lower sea level by ~2 m. The volumes of the lake basins and groundwater combine to yield a water storage potential equivalent to 4–8 m of sea level.

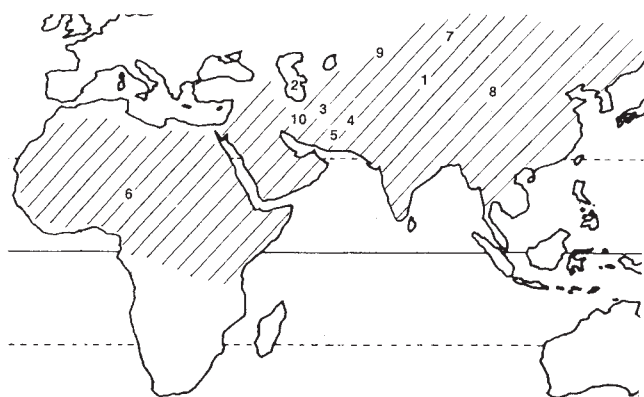


FIG. 1 Map of North Africa and South Asia. The hatched area is the region of increased monsoonal precipitation roughly 9,000 years ago as indicated by climate models⁸⁻¹⁰, lake level histories in Africa and Asia^{11,12}, and marine cores influenced by the Indus¹³, Nile¹⁴, and Niger¹⁵ rivers. This area of increased monsoonal precipitation extended over about 4×10^7 km², more than 25% of the Earth's continental area. The definition of the monsoon used includes deflections of the inter-tropical convergence zone away from the Equator and associated summertime tropical and subtropical rainfall either with or without discrete changes in wind direction. The 10 basins whose volume we determined are numbered as in Table 1. The volume of these basins and associated groundwater suggests a considerable potential for monsoonally driven sea-level change.

relationship in which periodic storage of water in lakes and aquifers led to sea-level fluctuation.

Models of eustatically produced nearshore carbonates assume rates of carbonate deposition and infer subsidence rates. Such models suggest that 10 m relative changes in sea level would optimize the production of the Lofer and Latemar cycles³⁻⁵. Isostatic loading contributes to relative sea-level change³² and reduces the required eustatic change to ~7 m, within the range of water storage suggested by Holocene monsoonal fluctuation (Table 1), and well within the lacustrine and groundwater storage potential associated with the stronger Triassic monsoon.

An ice-sheet explanation for the sea-level changes implied by the modelling of Triassic carbonates would require the expansion and melting of an ice cap larger than that of Greenland every 20 kyr (ref. 33). There is no evidence for a Triassic ice sheet even a fraction of the size of the Greenland ice cap. In addition, fossil floras containing ferns and cycads are known from high palaeo-latitudes in the Triassic, indicating that climates were equable near the poles³⁴. There is thus considerable evidence that cyclic lacustrine and groundwater storage^{7,25}, rather than changes in ice volume made the main contribution to Late Triassic sea-level fluctuation.

Although the evidence for a Milankovitch-forced monsoonal influence on sea level is most compelling in the Triassic, there is evidence for sea-level fluctuation contemporaneous with extensive lacustrine water storage in the Jurassic, Cretaceous and

Devonian. These periods also lack strong evidence of continental glaciation⁶. Small-scale sea-level fluctuations are evident in the lower Jurassic of Northern Italy³⁵, and lacustrine deposition continues into the lower Jurassic in the Newark Supergroup^{7,29}. The incipient opening of the South Atlantic produced a Lower Cretaceous salt basin north of the Walvis ridge and graben structures containing lacustrine deposits in South America and Africa³⁶. The Early Devonian was a time of fluctuating sea

level³⁷, continental red beds suggestive of aridity and climate fluctuation³⁸, and lake sediments containing Milankovitch cycles³⁹.

Climate-driven fluctuations in lake and groundwater volume could contribute only a small fraction of the sea-level changes observed in the Quaternary; however, in geological periods lacking continental ice sheets, lacustrine and groundwater storage may dominate the Milankovitch eustatic signal. □

Received 1 October 1992; accepted 11 January 1993.

1. Milankovitch, M. K. *Serb. Akad. Beogr. Sec. Publ.* **132**, (1941).
2. Hays, J. D., Imbrie, J. & Shackleton, N. J. *Science* **194**, 1121–1132 (1976).
3. Goldammer, R. K., Dunn, P. A. & Hardie, L. A. *Am. J. Sci.* **287**, 853–892 (1987).
4. Goldammer, R. K., Dunn, P. A. & Hardie, L. A. *Geol. Soc. Am. Bull.* **102**, 535–562 (1990).
5. Hinnov, L. A. & Goldammer, R. K. *J. sedim. Petrol.* **61**, 1173–1193 (1991).
6. Hambrey, M. J. & Harland, W. B. *Earth's Prepleistocene Glacial Record* (Cambridge Univ. Press, Cambridge, 1981).
7. Olsen, P. E. *Science* **234**, 789–912 (1986).
8. Kutzbach, J. E. *Science* **214**, 59–61 (1981).
9. Kutzbach, J. E. & Street-Perrott, F. A. *Nature* **317**, 130–134 (1985).
10. COHMAP *Science* **241**, 1043–1052 (1988).
11. Street, F. A. & Grove, A. T. *Quat. Res.* **12**, 83–118 (1979).
12. Lister, G. et al. *Palaeogeogr. Palaeoclim. Palaeoecol.* **84**, 141–162 (1991).
13. Van Campo, E. *Quat. Res.* **26**, 376–388 (1986).
14. Rossignol-Strick, M. *Nature* **304**, 46–49 (1983).
15. Pastoret, L., Chamey, H., Delibrias, G., Duplessy, J. C. & Thiede, J. *Oceanol. Acta* **1**, 217–232 (1978).
16. Kröpelin, S. & Soulié-Marsche, I. *Quat. Res.* **36**, 210–223 (1991).
17. Haynes, C. V. et al. *Quat. Sci. Rev.* **8**, 109–136 (1989).
18. Gasse, F. et al. *Palaeogeogr. Palaeoclim. Palaeoecol.* **60**, 1–46 (1987).
19. Ritchie, J. C., Eyles, C. H. & Haynes, C. V. *Nature* **314**, 352–355 (1985).
20. Grunnert, J., Baumhauer, R. & Völkel, J. *J. Afr. Earth Sci.* **12**, 133–146 (1991).
21. Tardy, Y., N'Kounkou, R. & Probst, J. *Am. J. Sci.* **289**, 455–483 (1989).

22. Street-Perrott, F. A. & Perrott, R. A. *Nature* **343**, 607–612 (1990).
23. Lautenschlager, M. *Geol. Rundschau* **80**, 513–534 (1991).
24. Dawson, A. G. *Ice Age Earth Late Quaternary Geology and Climate* (Routledge, London, 1992).
25. Dubiel, R. F., Parrish, J. T., Parrish, J. M. & Good, S. C. *Paleosols* **6**, 347–370 (1991).
26. Kutzbach, J. E. & Gallimore, R. G. *J. geophys. Res.* **94**, 3341–3357 (1989).
27. Holzer, W. T., Clement, G. P., Jansa, L. F. & Wade, J. A. in *Triassic and Jurassic Rifting* (ed. Manspeizer, W.) 523–553 (Elsevier, Amsterdam, 1988).
28. Ziegler, P. A. in *Triassic and Jurassic Rifting* (ed. Manspeizer, W.) 711–749 (Elsevier, Amsterdam, 1988).
29. Smoot, J. P. *Palaeogeogr. Palaeoclim. Palaeoecol.* **84**, 369–423 (1991).
30. Brown, R. H. *Am. Assoc. Petr. Geol. Bull.* **64**, 988–1003 (1980).
31. Gordon, W. A. *J. Geol.* **83**, 671 (1975).
32. Pitman, W. C. *Geol. Soc. Am. Bull.* **89**, 1389–1403 (1978).
33. Letréguilly, A., Huybrechts, R. & Reeh, N. *J. Glaciol.* **37**, 149–157 (1991).
34. Ash, S. R. & Basinger, J. F. *Geol. Surv. Can. Bull.* **412**, 101–131 (1991).
35. Bossellini, A. in *Controls on Carbonate Platform and Basin Development* (eds Crevello, P. D., Wilson, J. L., Sarg, J. F. & Read, J. F.) 3–14 (SEPM Spec. Publ. no. 44, 1989).
36. Uchupi, E. *J. Afr. Earth Sci.* **8**, 143–164 (1989).
37. Anderson, E. J., Goodwin, P. W. & Sobieski, T. H. *Geology* **12**, 120–123 (1984).
38. Ziegler, P. A. *Evolution of Laurasia* (Kluwer, Dordrecht, 1989).
39. Rogers, D. A. & Astin, T. R. *Spec. Publ. int. Ass. Sediment.* **13**, 199–221 (1991).

ACKNOWLEDGEMENTS. We thank J. B. Bennington, P. E. Olsen, J. F. Read, N. H. Shubin & A. M. Ziegler for discussions and J. E. Kutzbach for comments.

The temperatures and oxygen-isotope composition of early Devonian oceans

Guoqiu Gao

Department of Geological Sciences, University of Texas at Austin, Texas 78712, USA

THE oxygen isotope composition of pre-Carboniferous (>360 Myr old) marine carbonates, cherts and phosphates has been reported^{1–8} to be lower than that of later (post-Devonian) samples. According to one explanation^{1,3}, this is a reflection of warmer (generally $\geq 40^\circ\text{C}$) pre-Carboniferous oceans relative to modern oceans. Alternatively, it has been proposed^{5,7} that the pre-Carboniferous oceans were depleted in ^{18}O (by 2‰ SMOW) relative to modern oceans. Here I report high $\delta^{18}\text{O}$ values (–1.9 to –2.9‰ PDB) in normal, shallow-marine limestones (and three brachiopod shells) from the Lower Devonian Haragan and Bois d'Arc formations in south-central Oklahoma. I interpret these values as near-primary, and therefore constraining the temperature and oxygen isotope composition of early Devonian sea water to $25 \pm 7^\circ\text{C}$ and 0 ± 1 ‰ SMOW respectively. In conjunction with similarly high $\delta^{18}\text{O}$ values obtained from older (Ordovician and Silurian) samples⁸, these results imply that the temperature and oxygen isotope composition of pre-Carboniferous oceans may, at least during some time intervals, have been similar to those of modern oceans.

Limestones were sampled from the Haragan and Bois d'Arc formations of the upper Hunton Group of Oklahoma. These formations have been dated to be Lochkovian (early Devonian) in age by brachiopod and conodont biostratigraphy^{9–11}. These two formations, less than 100 m thick, crop out locally in the Arbuckle mountains of south-central Oklahoma, and consist mainly of skeletal wackestones and lime mudstones. Samples were collected in the Arbuckle mountains at a quarry located in section 31, T1S and R2E of Oklahoma. In previous studies^{4–8}, brachiopod shells and marine cements were selected for oxygen isotope analysis to reconstruct the temperatures and oxygen isotopic composition of ancient oceans. Here I chose lime mud

matrix without visible fossils and cements for isotopic and elemental analysis. Sampling sites were located on thin, polished slabs under both binocular and cathodoluminescence microscopes. Except for three dolomite-containing samples, sample powders were obtained from homogeneous, nonluminescent lime mud matrix by using a bench-top drill with dental drill bits. Three brachiopod shells were later analysed to compare to the carbon and oxygen isotopic data of the lime mud matrix.

The $\delta^{13}\text{C}$ and $\delta^{18}\text{O}$ values of the Haragan and Bois d'Arc limestones range from +1.1 to +2.6‰ (PDB) and –1.9 to –4.0‰ (PDB), respectively (Table 1). The $\delta^{13}\text{C}$ values (+1.8 to +2.0‰) of three brachiopod shells fall within the $\delta^{13}\text{C}$ range of the limestones, and the $\delta^{18}\text{O}$ values (–2.1 to –2.6‰) of the shells are similar to those (–1.9 to –2.9‰) of most limestones (22 of 25 analysed samples). The similarity suggests that the isotope data for the lime mud matrix are reliable, because brachiopod shells, composed of relatively stable low-Mg calcite, are considered by many^{4–6,8} to be the ideal carbonate component for reconstruction of the temperatures and oxygen isotopic composition of ancient oceans. Although the $\delta^{13}\text{C}$ values of the Haragan and Bois d'Arc limestones are similar to Lower Devonian carbonates elsewhere, the $\delta^{18}\text{O}$ values (–1.9 to –2.9‰) of most limestones (including three brachiopod shells, hereafter) are the highest $\delta^{18}\text{O}$ values yet documented for lower Devonian limestones (Fig. 1)^{5–7,13,14}. In fact, such high $\delta^{18}\text{O}$ values are not common in pre-Mississippian carbonates^{4–8}.

There could be several reasons for the high $\delta^{18}\text{O}$ values of the Haragan and Bois d'Arc limestones, including mineralogical impurity (such as the presence of dolomite), unusual depositional settings, and post-depositional (diagenetic) alteration. Dolomite should be enriched in ^{18}O by 2–4‰ (PDB) relative to calcite if both minerals are formed under similar conditions¹⁵. Thus, a marine limestone with minor amounts of syndepositional marine dolomite should have a higher $\delta^{18}\text{O}$ value than a pure limestone. For the analysed Haragan and Bois d'Arc limestones, minor amounts (<5 mol %) of dolomite were observed petrographically and detected by X-ray diffraction in three samples (H2-2, H2-5 and H2-8; Table 1). These three samples have lower $\delta^{18}\text{O}$ values (–3.4 to –4.0‰, Fig. 1), as well as higher Fe and Mn concentrations (Table 1), than dolomite-free

samples, so the dolomite in these samples must have low $\delta^{18}\text{O}$ values and high Fe and Mn concentrations relative to limestone. Such chemical compositions suggest that the dolomite is not of syndepositional marine origin and that the dolomite probably formed under reducing, burial conditions. The high $\delta^{18}\text{O}$ values of the Haragan and Bois d'Arc limestones cannot be attributed to contamination by syndepositional marine dolomite.

Two marine depositional settings can result in carbonate sediments with high $\delta^{18}\text{O}$ values relative to normal, shallow marine deposits. These are deep-water (low temperature) and restricted evaporative (^{18}O -enriched sea water) settings. Within the Haragan and Bois d'Arc formations, there is no sedimentological evidence for either a deep-water (>200 m) setting or an evaporative setting¹⁶. In the palaeontologic record, the Haragan and Bois d'Arc formations contain a rich, large-shelly fauna dominated by brachiopods^{9,10}, with many ostracodes, trilobites and bryozoans. This, together with lithostratigraphic evidence, indicates that the Haragan and Bois d'Arc formations were deposited in a normal, shallow marine (subtidal) environment¹⁶. The high $\delta^{18}\text{O}$ values of the limestones studied are therefore unlikely to be due to unusual environmental causes.

Kastner *et al.*¹⁷ documented a case in which diagenetic alteration resulted in high $\delta^{18}\text{O}$ values of the altered carbonate sediments. They found that recrystallization of calcareous ooze to chalk and limestone by normal sea water in low-temperature deep-sea conditions was accompanied by an increase in $\delta^{18}\text{O}$ values. But the deep-sea setting cannot be applied to the Haragan and Bois d'Arc limestones, which are shallow marine in origin. In theory, alteration of shallow-water limestones by ^{18}O -enriched water at low temperature (that is, at near-surface conditions) could result in limestones with high $\delta^{18}\text{O}$ values. The ^{18}O -enriched water could be either evaporative sea water or basinal brines. For the Haragan and Bois d'Arc limestones, diagenetic ^{18}O enrichment by late evaporative sea water is unlikely, as these limestones were not followed by deposition of evaporative sediments¹⁶. Basinal brines are usually characterized by high (radiogenic) $^{87}\text{Sr}/^{86}\text{Sr}$ ratios as a result of water-silicate interactions in sedimentary basins¹⁸, and this should be

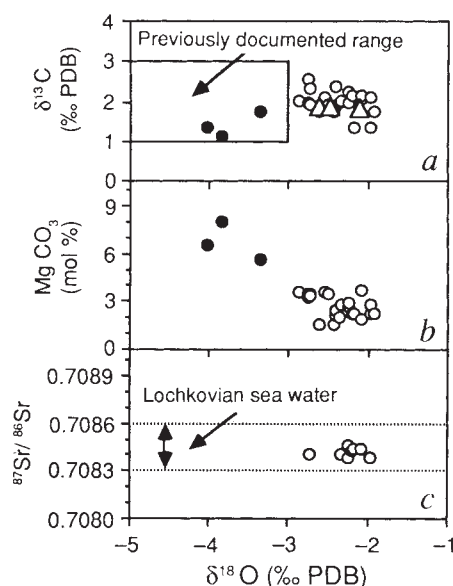


FIG. 1 Plots of $\delta^{18}\text{O}$ against $\delta^{13}\text{C}$, MgCO_3 and $^{87}\text{Sr}/^{86}\text{Sr}$ for the Haragan and Bois d'Arc samples (Δ : brachiopod shells; $\circ$: dolomite-free lime mud samples; $\bullet$: dolomite-containing samples). These plots show that (a) most of these samples have higher $\delta^{18}\text{O}$ values than previously documented samples^{5-7,13,14}, (b) three dolomite-containing samples have lower $\delta^{18}\text{O}$ values and higher MgCO_3 concentrations than other dolomite-free samples, and (c) the $^{87}\text{Sr}/^{86}\text{Sr}$ ratios of lime mud samples fall in the $^{87}\text{Sr}/^{86}\text{Sr}$ range of coeval sea water¹⁹.

TABLE 1 Chemical composition of Haragan-Bois d'Arc limestone

Sample	$\delta^{13}\text{C}$ (PDB)	$\delta^{18}\text{O}$ (PDB)	$^{87}\text{Sr}/^{86}\text{Sr}$ $\pm 2\sigma^*$	CaCO_3 (mol%)	MgCO_3 (mol%)	Fe (p.p.m.)	Mn (p.p.m.)	Sr (p.p.m.)
Lime mud matrix								
H2-1	+1.93	-2.74	0.708407 $\pm$ 11	96.111	3.539	1.692	87.7	278.1
duplicate				96.154	3.498	1.686	87.7	279.3
H2-2 [†]	+1.35	-4.03		92.806	6.556	3.247	90.6	398.5
H2-3	+2.11	-2.56		96.065	3.602	1.450	87.1	243.4
H2-4	+2.02	-2.35	0.708401 $\pm$ 11	96.924	2.786	1.398	94.7	207.6
duplicate	+2.08	-2.41						
H2-5 [‡]	+1.11	-3.85		91.023	7.950	5.524	99.7	355.7
H2-6	+1.76	-2.43		98.274	1.532	878	88.6	182.8
H2-7	-2.15	-2.20	0.708430 $\pm$ 11	97.553	2.255	867	85.6	188.1
H2-8 [‡]	+1.76	-3.36		93.788	5.611	3.077	95.3	324.3
H2-9	+2.03	-2.87		95.985	3.626	1.921	89.9	266.2
H2-10	+2.17	-2.08	0.708441 $\pm$ 12	97.902	1.902	565	78.3	178.2
H2-11	+2.33	-2.74		96.396	3.338	1.257	84.9	232.1
H2-12	+1.75	-2.63		98.337	1.550	434	83.2	174.0
H2-13	+1.37	-1.98	0.708387 $\pm$ 08	97.037	2.800	690	88.8	208.1
duplicate	+1.46	-1.92						
H2-14	-2.09	-1.98		97.601	2.276	487	78.7	191.9
duplicate				97.575	2.305	477	79.8	189.6
H2-15	+1.75	-1.93		97.633	2.231	557	88.2	181.0
H2-16	+1.98	-2.26	0.708449 $\pm$ 13	97.209	2.663	508	86.5	188.1
duplicate			0.708466 $\pm$ 09					
H2-17	+1.98	-2.76		96.165	3.529	1.442	94.0	282.8
H2-18	+2.00	-2.36		97.913	1.967	457	81.6	204.3
H2-19	+1.87	-2.42		97.715	2.136	601	95.7	213.9
H2-20	+1.35	-2.18	0.708447 $\pm$ 15	97.650	2.223	486	110.8	179.1
duplicate	+1.26	-2.24	0.708437 $\pm$ 10					
HT1	+2.06	-2.09		95.995	3.740	1.254	90.2	215.0
HT2	+1.94	-2.50		96.240	3.460	1.425	91.4	267.3
HT3	+2.36	-2.41		97.302	2.466	1.058	75.0	257.7
HT4	+2.24	-2.25	0.708381 $\pm$ 11	96.929	2.895	757	84.6	230.0
HT5	+2.55	-2.76		96.569	3.259	710	82.3	268.7
duplicate				96.480	3.346	715	83.9	269.9
Brachiopod shells								
Bra1	+1.89	-2.47						
Bra2	+1.85	-2.12						
Bra3	+1.90	-2.63						

Methods and procedures for isotopic and elemental analyses were similar to those described in ref. 12.

* $^{87}\text{Sr}/^{86}\text{Sr}$ ratios reported here are normalized relative to NBS 987=0.71014 and $2\sigma \times 10^{-6}$ is the uncertainty.

[‡] Samples with minor amount of dolomite

reflected in altered limestones. Yet the $^{87}\text{Sr}/^{86}\text{Sr}$ ratios of the studied limestones fall tightly in the $^{87}\text{Sr}/^{86}\text{Sr}$ range of early Devonian sea water¹⁹ (Fig. 1). This indicates that they were not modified by basinal brines.

Diagenetic alteration almost always results in ^{18}O depletion of carbonates, as alteration commonly occurs in the presence of ^{18}O -depleted meteoric water or under burial (elevated temperature) conditions^{15,20}. As well as ^{18}O depletion, altered limestone usually displays textural recrystallization and elemental change (Sr depletion and Fe and Mn enrichment)²⁰. Within the Haragan and Bois d'Arc limestones, the original micrite has largely recrystallized to microspar, as ascertained by petrographic examination. The limestones now have lower Sr (<300 p.p.m.) and higher Fe (430–1,920 p.p.m.) and Mn (75–110 p.p.m.) concentrations relative to unaltered modern marine sediments (Sr 200–9,000 p.p.m.; Fe and Mn both <40 p.p.m.)^{20,21}. Their texture and original elemental chemistry have been altered during diagenesis from whatever precursor minerals were present. If such alteration occurred in the presence of meteoric water or under burial conditions, the most altered limestone should have the lowest $\delta^{18}\text{O}$ value and Sr content, as well as the highest Fe and Mn concentrations. Such a chemical variation is not observed (Fig. 2). I propose that the textural and elemental alteration of the Haragan and Bois d'Arc limestones probably occurred early in a marine phreatic environment. Early marine diagenesis should not greatly affect the primary $\delta^{18}\text{O}$ signature of these limestones.

As the high $\delta^{18}\text{O}$ values (−1.9 to −2.9‰ PDB) of the Haragan and Bois d'Arc limestones are unlikely to have resulted from

the presence of syndepositional marine dolomite and diagenetic alteration, I interpret them to be near-primary signals. Furthermore, considering the fact that the Haragan and Bois d'Arc limestones were deposited in a low-latitude ($<35^\circ$), shallow, epicontinental sea^{22,23}, I suggest that they represent the best-preserved, near-primary $\delta^{18}\text{O}$ values for low-latitude, shallow marine limestones elsewhere. The values are considerably higher than those ($\delta^{18}\text{O} \leq -3.0\%$) of the previously documented, best-preserved Lower Devonian brachiopods and marine calcitic cements^{5-7,14}.

The issue of whether the oxygen isotopic composition of sea water has changed in geological time is much debated. One group proposes that sea water $\delta^{18}\text{O}$ has been relatively invariant, because theoretical calculations indicate that hydrothermal alteration (sea water/basalt reaction) at mid-ocean ridges and continental weathering should buffer sea water $\delta^{18}\text{O}$ to $0 \pm 1\%$ (SMOW)²⁴⁻²⁶. Assuming a constant $\delta^{18}\text{O}$ for sea water, some researchers suggest that the low $\delta^{18}\text{O}$ values of pre-Mississippian carbonates, cherts and phosphates relative to those of post-Devonian counterparts^{1,3} arose because pre-Mississippian oceans generally had high temperatures ranging from 40 to 75°C (ref. 3). Hot world oceans, however, appear to be inconsistent both with large-scale glaciation during the pre-Mississippian time and with the upper temperature limit (about 38°C) of living marine communities⁵. Therefore, others have postulated^{5,7} that seawater $\delta^{18}\text{O}$ must have changed with time, with early Devonian $\delta^{18}\text{O}$ values being less than -2% (SMOW). But the near-primary $\delta^{18}\text{O}$ values from -1.9 to -2.9% (PDB) for low-latitude, shallow marine limestones found here constrain the temperatures and $\delta^{18}\text{O}$ values of early Devonian sea water to $25 \pm 7^\circ\text{C}$ and $0 \pm 1\%$ (SMOW), respectively, according to the relationship $10^3 \ln \alpha_{\text{calcite-water}} = 2.78 \times 10^6 T^{-2} - 2.89$ (where α is the oxygen isotope fractionation factor between calcite and water, and T is temperature in kelvin)²⁷. This indicates that the temperatures and oxygen isotopic composition of early Devonian oceans were similar to those of modern oceans.

Documentation of primary $\delta^{18}\text{O}$ values for unaltered marine carbonates, cherts and phosphates could potentially provide a solution to the controversy about the temperatures and oxygen isotopic composition of pre-Mississippian oceans. The primary record may be a band of 2 to 4% (ref. 8), because of variations in the temperatures and $\delta^{18}\text{O}$ values of the worldwide oceans. However, the highest $\delta^{18}\text{O}$ values of marine carbonates, cherts and phosphates should provide the best constraints, because

diagenetic alteration almost always lowers the $\delta^{18}\text{O}$ values of sedimentary minerals and rock^{15,20}. Recently, several studies have documented near-primary, high $\delta^{18}\text{O}$ values for low-latitude marine limestones and brachiopods of Early Palaeozoic age (specifically, Early Ordovician: -3.1% ; Late Ordovician: -1.4 to -2.4% ; and Late Silurian: -1.9% (refs 8, 28, 29). These values, together with the results reported here, suggest that pre-Carboniferous oceans may have been (at least during some intervals) similar in temperature and $\delta^{18}\text{O}$ to those of today. Additional oxygen isotope studies of well preserved carbonates, cherts and phosphates will provide further tests of this assertion. $\square$

Received 17 August 1992; accepted 5 January 1993.

1. Knauth, L. P. & Epstein, S. *Geochim. cosmochim. Acta* **40**, 1095-1108 (1976).
2. Shemesh, A. et al. *Earth planet. Sci. Lett.* **64**, 405-416 (1983).
3. Karhu, J. & Epstein, S. *Geochim. cosmochim. Acta* **50**, 1745-1756 (1986).
4. Popp, B. N. et al. *Geol. Soc. Am. Bull.* **97**, 1262-1269 (1986).
5. Veizer, J. et al. *Geochim. cosmochim. Acta* **50**, 1679-1696 (1986).
6. Hudson, J. D. & Anderson, T. F. *Trans. R. Soc. Edinb.* **80**, 183-192 (1989).
7. Lohmann, K. C. & Walker, J. C. G. *Geophys. Res. Lett.* **16**, 319-322 (1989).
8. Wadleigh, M. A. & Veizer, J. *Geochim. cosmochim. Acta* **56**, 431-443 (1992).
9. Amsden, T. W. *Oklahoma Geol. Surv. Bull.* **78**, 9-144 (1958a).
10. Amsden, T. W. *Oklahoma Geol. Surv. Bull.* **82**, 1-110 (1958b).
11. Barrick, J. E. et al. *Oklahoma Geol. Surv. Guidebook* **27**, 55-62 (1990).
12. Gao, G. *Geochim. cosmochim. Acta* **54**, 1979-1989 (1990).
13. Rush, P. F. & Chafetz, H. S. *J. sedim. Petrol.* **60**, 968-981 (1990).
14. Lavoie, D. *Geol. Soc. Am. Abstr. Progr.* **24**, A37 (1992).
15. Land, L. S. in *Concepts and Models of Diagenitization* (eds Zenger D. H. et al.). *Soc. Econ. Palaeont. Mineral. Spec. Publ.* **28**, 87-110 (1980).
16. Amsden, T. W. *Am. Assoc. Petrol. Geol. Bull.* **46**, 1502-1519 (1962).
17. Kastner, M. et al. *Nature* **321**, 158-161 (1986).
18. Chaudhuri, S. et al. *Geochim. cosmochim. Acta* **51**, 45-53 (1987).
19. Burke, W. H. et al. *Geology* **10**, 516-519 (1982).
20. Veizer, J. *Soc. Econ. Palaeont. Mineral. Short Course* **10**, 3-1 to 3-100 (1983).
21. Carpenter, S. J. et al. *Geochim. cosmochim. Acta* **55**, 1991-2010 (1991).
22. Heckel, P. H. & Witzke, B. J. *Palaeont. Assoc. London spec. Pap. Palaeontol.* **23**, 99-123 (1979).
23. Johnson, K. S. et al. in *The Geology of America*, vol. D-2 *Sedimentary Cover-North American Craton* (ed. Sloss, L. L.) 307-359 (Geol. Soc. Am., Boulder, Colorado, 1988).
24. Muehlenbachs, K. & Clayton, R. N. *J. geophys. Res.* **81**, 4365-4369 (1976).
25. Gregory, R. T. & Taylor, H. P. *J. geophys. Res.* **86**, 2737-2755 (1981).
26. Holland, H. D. *The Chemical Evolution of the Atmosphere and Oceans* (Princeton Univ. Press, 1984).
27. Friedman, I. & O'Neill, J. R. *U.S. Geol. Surv. Prof. Pap.* **440K** (1977).
28. Marshall, J. D. & Middleton, P. D. *J. geol. Soc. London* **147**, 1-4 (1990).
29. Gao, G. & Land, L. S. *Geochim. cosmochim. Acta* **55**, 2911-2920 (1991).

ACKNOWLEDGEMENTS. I thank James Burdett and Lynton Land for constructive reviews. The Geology Foundation, University of Texas at Austin, defrayed certain publication costs.

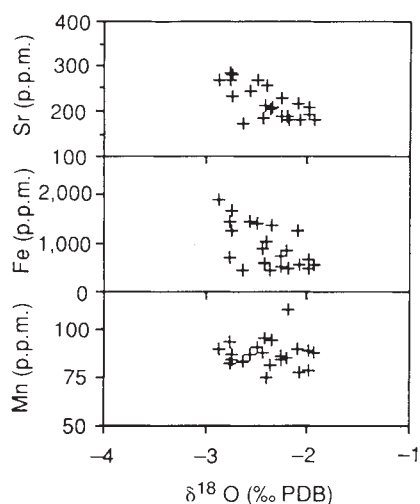


FIG. 2 Plots of $\delta^{18}\text{O}$ against Sr, Fe and Mn for the Haragan and Bois d'Arc lime mud limestones (three dolomite-containing samples are excluded). These plots show that limestones with lower $\delta^{18}\text{O}$ values are not simultaneously characterized by lower Sr and higher Fe and Mn concentrations.

The 1992 Nicaragua earthquake: a slow tsunami earthquake associated with subducted sediments

Hiroo Kanamori* & Masayuki Kikuchi†

* Seismological Laboratory, California Institute of Technology, Pasadena, California 91125, USA

† Department of Physics, Yokohama City University, Yokohama 236, Japan

THE 1992 Nicaragua earthquake was a 'tsunami earthquake'; that is, it generated tsunamis¹ disproportionately large for its surface-wave magnitude, M_s . The moment magnitude, M_w , determined from long-period (~ 250 -s) surface waves², was 7.6, significantly larger than the 20-s M_s of 7; this M_s - M_w disparity is also characteristic of tsunami earthquakes^{3,4}. The Nicaragua earthquake is the first tsunami earthquake to be captured by modern broadband seismic networks, allowing us to present here seismograms of sufficiently high quality to make inferences about the rupture mechanisms. We conclude that the Nicaragua earthquake was a slow thrust earthquake which occurred on the subduction interface between the Cocos and North American plates, and because of the absence of sediments on the trench floor offshore of Nicaragua, the slip propagated up-dip all the way to the ocean

bottom, exciting large tsunamis. The occurrence of slip on a plate interface filled with soft subducted sediments caused the rupture process to be slower than in ordinary subduction-zone thrust earthquakes. Our results reinforce the idea that tsunami warning systems using long-period (≥ 100 -s) waves⁵⁻⁷ are necessary to reduce the hazard from this type of earthquake.

Figure 1 shows the mainshock epicentre, the aftershock area, the mechanism determined from long-period surface waves and the source time function determined from long-period body waves. The surface-wave magnitude, M_s , determined with data from 11 IRIS stations is 7. The epicentre is near a previously identified seismic gap⁸. The mechanism is thrust which is consistent with subduction of the Cocos plate beneath Nicaragua. A seismic moment, M_0 , of 3.7×10^{27} dyn cm ($M_w = 7.6$) was obtained. The seismic moment could be considerably larger or smaller than this value if the dip angle, which is not well determined, was reduced or increased, respectively. This solution is similar to the Harvard centroid moment tensor (CMT) solution¹. The difference is partly due to the difference in the dip angle and the period of the wave used. The long-period centroid location is closer to the trench than the epicentre, suggesting that the rupture propagation was up-dip toward the trench.

The anomalous character was readily identified on the high-quality seismograms. For example, Fig. 2 shows a long-period wave, indicated by an arrow, on the displacement seismogram recorded at the Pasadena TERRAscope station. For ordinary earthquakes, this long-period wave is buried under much larger short-period waves and is not visible. The appearance of this wave is a direct manifestation of deficiency in short-period energy. Figure 3 compares long-period Rayleigh waves from the Nicaragua earthquake recorded at station KON (Norway) with those from the 28 June 1992, earthquake in Landers, California ($M_s = 7.5$, $M_w = 7.3$) recorded at station MAJO (Japan). The two seismograms are plotted to the same scale, so that the amplitude can be compared directly. These stations are located at about the same distance from their respective epicentres and in the azimuth where the Rayleigh-wave radiation is maximum. The difference in the spectrum is evident. The very-long-period Rayleigh waves (up to a period of 500 s) for the Nicaragua earthquake are much larger than those for the Landers earthquake, for both R_1 and R_2 . In contrast, short-period surface waves are larger for the Landers earthquake than for the Nicaragua earthquake. This observation demonstrates the dis-

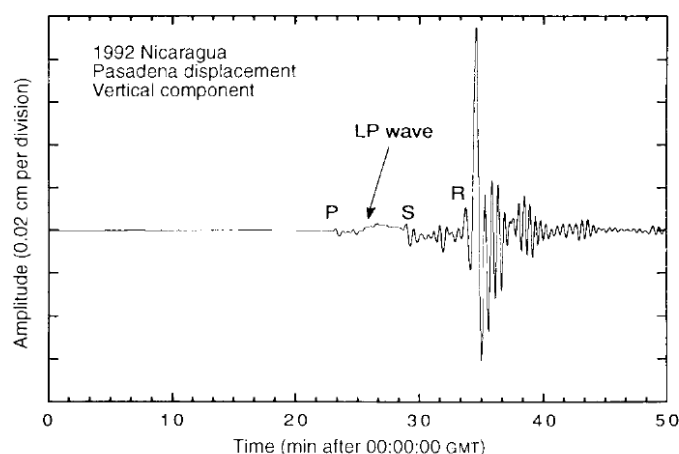


FIG. 2 The displacement seismogram recorded at the Pasadena TERRAscope station. Note the large long-period wave indicated by an arrow.

parity between M_s and M_w , and indicates that the Nicaragua earthquake is a slow earthquake.

We examined seismic excitation at periods longer than 200 s by comparing the observed normal-mode amplitude with that computed for the source mechanism shown in Fig. 1. No evidence for enhanced seismic excitation was found beyond a period of 200 s. Thus, the Nicaragua earthquake is anomalous over a period range of up to 200 s.

The Nicaragua earthquake has very long source duration. The centroid origin time of the Nicaragua earthquake is ~ 50 s after the origin time determined by the National Earthquake Information Center (NEIC) using short-period waves. Inversion of body waves, shown in Fig. 1, also indicates a source duration of 80–100 s although the details of the source function could not be resolved. These observations suggest that the overall source duration is ~ 100 s. The broadband response of modern seismic networks is essential for determining source duration.

Because submarine slumping has been suggested as a cause of some tsunamis⁹⁻¹¹, we tried a slump model to interpret the Nicaragua earthquake data. We used a horizontal single force, instead of a moment tensor, as a kinematic representation of slumping, but a single-force model simply could not fit the data. Thus we conclude that the primary source of the Nicaragua earthquake is a slow thrust fault on the subduction interface, although some slumping could have accompanied the earthquake. Whether the size of the Nicaragua earthquake determined from long-period waves is large enough to explain the observed

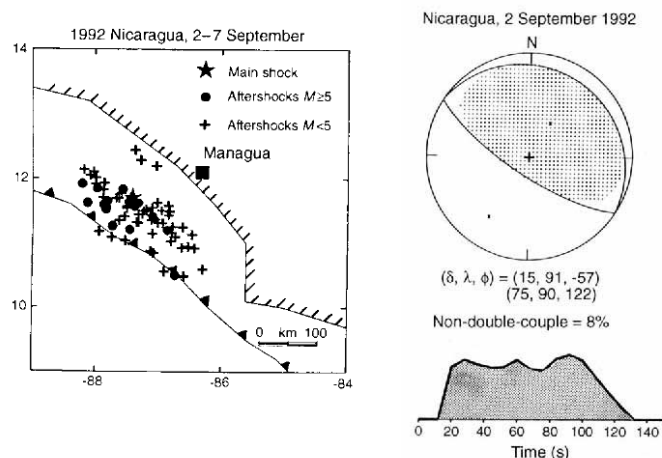


FIG. 1 Location of the main shock and the aftershocks of the 2 September 1992 Nicaragua earthquake (left), the mechanism determined with long-period surface waves from 11 stations (upper right), and the source time function determined with long-period body waves from six stations (lower right). The non-double-couple component of 8% is commonly found in this type of inversion, and should be considered insignificant. The body-wave solution was obtained using the first 150 s of P and SH waves from six stations, with the mechanism constrained to be the same as that determined with surface waves.

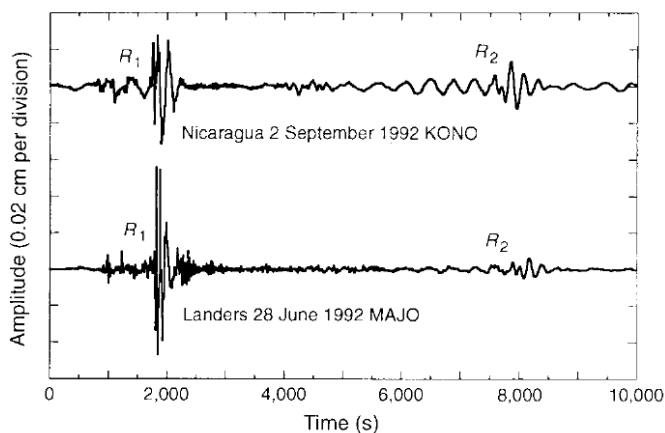
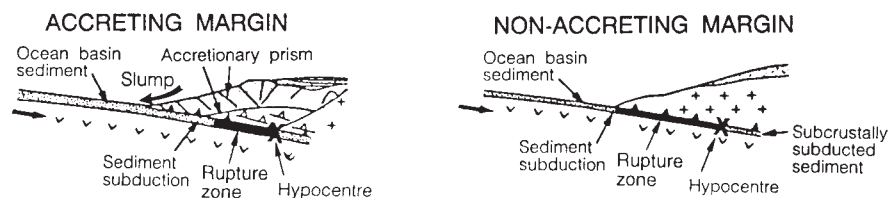


FIG. 3 Comparison of the long-period surface waves (displacement) from the Nicaragua earthquake ($M_s = 7.0$, $M_w = 7.6$) with those from the Landers earthquake ($M_s = 7.5$, $M_w = 7.3$). Note the large long-period (up to 500 s) waves seen on the record of the Nicaragua earthquake, for both R_1 and R_2 .

FIG. 4 Schematic diagrams showing the two types, accreting and non-accreting, of plate margin²⁰. The Nicaragua earthquake occurred on a non-accreting margin where the sediments are completely subducted. The slow rupture extending all the way to the trench floor causes tsunamis. In accreting margins, thrust earthquakes may not rupture to the trench floor, but occasional sediment slumping on the accretionary prism may cause very large tsunami earthquakes.



tsunami is unresolved at present. The excitation and propagation of tsunamis are strongly affected by the coastal topography, focusing and defocusing of tsunami energy, and the details of the sea-floor deformation.

Among tsunami earthquakes, the 1896 Sanriku, Japan, and the 1946 Unimak Islands (Aleutian Islands) earthquakes are most anomalous³. The Sanriku earthquake was felt only moderately along the Sanriku coast, but very large tsunamis hit the coast ~30 min later and drowned more than 22,000 people; presumably, because of the absence of strong shaking, they did not expect large tsunamis. The magnitudes of these earthquakes determined at short period are only 6.8 to 7.5 (refs 3, 12), but they excited some of the largest tsunamis in history. The source spectral amplitude of these events increases beyond 200 s (ref. 3). Compared with these earthquakes the Nicaragua earthquake is not very anomalous; it is comparable to other tsunami earthquakes such as the 1960 Peru^{13,14}, 1963 Kurile Island (October 20)^{14,15} and 1975 Kurile Island earthquake^{14,15}.

Several mechanisms have been proposed for tsunami earthquakes. When tsunami earthquakes were first recognized, slow slip on the subduction interface was intuitively thought to be the cause of their deficiency in short-period energy³. It was later suggested that a secondary steeply dipping fault in the accretionary prism was responsible for some tsunami earthquakes¹⁵. Vertical displacement on a steep fault is considered more favourable for tsunami excitation. Theoretically, a shallow vertical dip-slip fault does not excite seismic waves efficiently but it generates tsunamis well; such a fault could thus be the source of tsunami earthquakes¹⁶. The moderate tsunami earthquakes 1960 Peru, 1963 Kurile Island and 1975 Kurile Island were all caused by slow slip on a basal decollement in the accretionary prism¹⁴. As shown above, the 1992 Nicaragua earthquake is similar to these events.

In contrast, the mechanism of the very anomalous tsunami earthquakes such as the 1896 Sanriku and 1946 Unimak Island earthquakes is unknown at present. A slump model has been suggested for these events¹⁰, but seismic data are limited both in quality and quantity, and cannot resolve their mechanism. In fact, a fault model is preferred for the 1946 Unimak Island earthquake¹⁷. In the Sanriku region, and along the Aleutian Islands, large-scale slump features are seen on the sea floor¹⁸. A large volume of sediments near the trench may occasionally collapse in slumping and cause large tsunamis. During the 1989 earthquake in Loma Prieta, California, tsunamis were excited by a slumping caused by shaking¹⁹.

Unlike the trenches along the Sanriku coast or the Aleutian Islands, the mid-American trench off Nicaragua does not have a large volume of sediments or a well-developed accretionary prism²⁰. This is surprising, because most tsunami earthquakes were previously thought to be associated with sediments in the accretionary prism. We suggest that the absence of sediments allowed the slip to propagate all the way to the ocean bottom. A large volume of soft sediments on the trench floor would prevent this.

Two types of tsunami earthquakes may then be suggested (Fig. 4). The first occurs in trenches with large amounts of sediment and an accretionary prism. Although the rupture of the individual thrust earthquake may not reach the ocean bottom, occasional slumping there may be the cause of large tsunami earthquakes. The second occurs in subduction zones

without large amounts of sediment. In these subduction zones, the sediments are completely subducted, and the plate interface is filled with soft sediments^{20,21}. The slip can extend to the surface, breaking through a relatively weak plate interface filled with sediments. The shallow depth and soft low-velocity sediments on the interface are responsible for efficient tsunami generation and slow rupture propagation, respectively. The 1960 Peru, 1963 and 1975 Kurile Islands, and 1992 Nicaragua earthquakes are of the second type, although the details in the rupture geometry (such as up-dip or rupture propagation) and the trench geomorphology may be different between these events. The truly large tsunami earthquakes (such as the 1896 Sanriku and 1946 Unimak Islands earthquakes) apparently belong to the first type, but direct evidence for this is not yet available. In future, however, the broadband capability of modern seismic networks will help seismologists to identify anomalous earthquakes such as slow earthquakes and tsunami earthquakes, and understand their relationship to the regional variation of subduction process.

Regardless of the mechanism, an important diagnostic feature of tsunami earthquakes is the difference between M_s and M_w . Unfortunately, most of the existing systems use an M_s threshold for issuing tsunami warnings. Long-period magnitudes such as M_w and mantle magnitude (a magnitude determined with mantle waves at a period of ~100 s)^{22,23} are more diagnostic of the tsunami potential of earthquakes^{3,4,22,23}. As is evident in Figs 2 and 3, with the use of M_w , the Nicaragua earthquake would have been easily diagnosed as a tsunami earthquake. In fact, its seismic moment was quickly estimated at long period at Papeete and reported to warning agencies within an hour of the earthquake origin time (E. Okal, personal communication). Several tsunami warning systems using long-period waves have been proposed for both teleseismic and local tsunami warning purposes³⁻⁷. Systematic use of such a system could reduce the hazards posed by tsunamis. □

Received 14 October 1992; accepted 4 January 1993.

- NEIC Quick Epicenter Determination, U.S. Department of the Interior (3 September 1992).
- Kanamori, H. *Nature* **271**, 411–414 (1977).
- Kanamori, H. *Phys. Earth planet. int.* **6**, 346–359 (1972).
- Abe, K. *J. geophys. Res.* **84**, 1561–1568 (1979).
- Kanamori, H. & Given, J. W. In *Tsunami—Their Science and Engineering* (eds Iida, K. & Iwasaki, T.) 37–49 (Terra Scientific, Tokyo, 1983).
- Kanamori, H. In *Int. Tsunami Symp.* (eds Murty, T. S. & Rapatz, W. J.) 30–37 (Inst. Ocean Sci. Sidney, British Columbia, 1985).
- Talandier, J. & Okal, E. A. *Bull. seism. Soc. Am.* **79**, 1177–1193 (1989).
- Harlow, D. H., White, R. A. & Cifuentes, I. L. *Science* **213**, 648–651 (1981).
- Gutenberg, B. *Bull. seism. Soc. Am.* **29**, 517–526 (1939).
- Kanamori, H. *Nondouble Couple Seismic Source* (abstr.) in *23rd Gen. Assembly Assoc. Seismology and the Physics of the Earth's Interior*, Aug. 19–30, 1985, 1–425 (Tokyo, 1985).
- Hasegawa, H. S. & Kanamori, H. *Seism. Soc. Am. Bull.* **77**, 1984–2004 (1987).
- Utsu, T. *Bull. Earthq. Res. Inst. Tokyo Univ.* **54**, 253–308 (1979).
- Pelayo, A. M. & Wiens, D. A. *Geophys. Res. Lett.* **17**, 661–664 (1990).
- Pelayo, A. M. & Wiens, D. A. *J. geophys. Res.* **97**, 15321–15337 (1992).
- Fukao, Y. *J. geophys. Res.* **84**, 2303–2314 (1979).
- Ward, S. N. *Bull. seism. Soc. Am.* **72**, 759–777 (1982).
- Pelayo, A. M. *Earthquake Source Parameter Inversion using Body and Surface Waves: Application to Tsunami Earthquakes and to Scotia Sea Seismotectonics* (Washington Univ., 1990).
- Cadet, J. P. et al. *Earth and Planetary Science Letters* **83**, 313–328 (1987).
- Ma, K.-F., Satake, K. & Kanamori, H. *Geophys. Res. Lett.* **18**, 637–640 (1991).
- von Huene, R. & Scholl, D. W. *Rev. Geophys.* **29**, 279–316 (1991).
- Hilde, T. W. C. *Tectonophysics* **99**, 381–397 (1983).
- Brune, J. N. & Engen, G. R. *Bull. seism. Soc. Am.* **59**, 923–933 (1969).
- Okal, E. A. & Talandier, J. *J. geophys. Res.* **94**, 4169–4193 (1989).

ACKNOWLEDGEMENTS. We benefited from discussions with G. Plafker, G. Suarez, D. Anderson, T. Heaton, S. Uyeda and T. Hilde. This research was conducted under the TERRAScope project supported by the L. K. Whittier and Arco Foundations.

Major role of ultraviolet-B in controlling bacterioplankton growth in the surface layer of the ocean

Gerhard J. Herndl, Gerald Müller-Niklas & Jürgen Frick

Institute of Zoology, University of Vienna, Althanstrasse 14, A-1090 Vienna, Austria

THERE is evidence that the potentially harmful solar ultraviolet-B (UV-B, 280–320 nm) radiation penetrates much deeper into the ocean's water column than previously thought^{1,2}. UV-B radiation is also responsible for photochemical degradation of refractory macromolecules into biologically labile organic compounds^{3,4}. It thus seems reasonable to assume that UV-B radiation might influence the cycling of organic matter in the sea, which is believed to be largely mediated by bacterioplankton⁵. Here we report that bacterioplankton activity in the surface layers of the oceans is suppressed by solar radiation by about 40% in the top 5 m of the water column in nearshore waters, whereas in oligotrophic open oceans suppression might be detectable to a depth of >10 m. Bacterioplankton from near-surface (0.5 m depth) waters of a highly stratified water column were as sensitive to surface UV-B radiation as subpycnocline bacteria, indicating no adaptive mechanisms against surface solar radiation in near-surface bacterioplankton consortia. Surface solar radiation levels also photochemically degrade bacterial extracellular enzymes. Thus elevated UV-B radiation due to the destruction of the stratospheric ozone layer might lead to reduced bacterial activity and accompanying increased concentration of labile dissolved organic matter in the surface layers of the ocean as bacterial uptake of this is retarded.

In evaluating the possible effects of natural solar radiation, we have to distinguish possible effects of short- and long-term exposure on the bacterial community. In a mixed water column, short-term exposure should predominate. We determined the influence of short-term (30 min) solar radiation on thymidine and leucine incorporation into bacterioplankton of the surface layer of the sea by incubating surface water at different UV-B radiation levels. Surface water samples were collected with

TABLE 1 Extracellular enzymatic activity of bacteria from subsurface waters

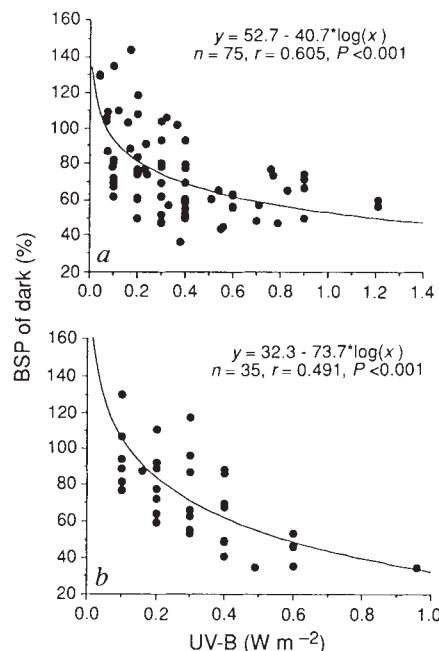
Enzyme	Exposure	Total EEA (EEA in % of initial activity)	Free EEA
β -Glucosidase	UV-B	8	ND
	4 h dark	31	ND
Lipase	UV-B	32	62
	4 h dark	124	
Leucine-aminopeptidase	UV-B	24	49
	4 h dark	62	

Extracellular enzymatic activity before and after a 4 h exposure of bacteria from subsurface waters (depth 0.5 m; incubation temperature, 26 °C) to natural sunlight (PAR+UV-A+UV-B) and subsequent 4 h incubation in dark. Incubations were done in 100 ml quartz flasks at around noon on a cloudless day. Extracellular enzymatic activity was measured using fluorogenic substrate analogues⁶. The respective substrate was added as a tracer yielding a final concentration of 2.5 μ M. The increase in fluorescence was measured in dark incubations over a 20 min period at 26 °C with a Jasco spectrofluorometer⁸; total EEA, total extracellular enzymatic activity; free EEA, extracellular enzymatic activity in particle-free 0.2 μ m-filtered sea water; all values are given in per cent of initial activity before the exposure to elevated UV-B radiation levels; ND, not detectable.

acid-cleaned Niskin bottles in the northern Adriatic Sea from depths ranging from 0.5 m to 2 m. Bacterial cell production as well as bacterial protein production was affected by UV-B radiation typically for surface layers (Fig. 1). Compared with bacterial activity in the dark, a 50% reduction in thymidine incorporation was observed at UV-B radiation levels of 1.3 $W m^{-2}$, whereas a 50% reduction in leucine incorporation was detectable at a radiation of about 0.6 $W m^{-2}$ (Fig. 1). Thus we conclude that, at least during short-term exposure, bacterial biomass production is more affected than cell division and hence cell production. The higher bacterial activity in the low UV-B radiation range (<0.2 $W m^{-2}$) compared to the dark activity might be caused by the release of easily metabolizable substances from the photosynthetic activity of phytoplankton.

In a stratified water column mixing is restricted and bacterioplankton in the surface waters may be exposed to high levels of UV-B during their entire life cycle, whereas bacterioplankton from the subpycnocline waters never experience noticeable levels of ultraviolet radiation. We addressed this problem by

FIG. 1 The dependence of bacterial secondary production (BSP) measured by [³H]thymidine (a) and [³H]leucine incorporation (b) at different UV-B radiation levels related to bacterial incorporation in the dark. Different radiation levels were obtained by using different mesh sizes to cover the quartz flasks. Bacterial incorporation was measured in 5 ml subsamples in triplicates with two formalin-killed blanks (2% final concentration) in sterile polystyrene test tubes (Falcon tubes) at *in situ* temperature. Incubation period was 0.5 h. Although thymidine incorporation measurements into bacterial DNA were meant to determine bacterial cell production^{9,10}, additional samples were incubated also with [³H]leucine to measure bacterial protein production (biomass production) in the same type of tubes¹¹. Wavelength scanning of these polystyrene tubes indicated that about 90% of the UV-B radiation penetrates the tubes. Occasionally, quartz flasks were also used; no significant differences were found between both incubation devices. UV-B penetration into the water column of the Gulf of Trieste (45° 33' N, 13° 35' E) and off Belize (16° 50' N, 88° 05' W) was measured using a semiconductor ultraviolet (G3614-01 Hamamatsu) sensor with a cutoff at the lower end of the spectrum at 260 nm and at 320 nm at the upper end of the spectrum with a peak response at 290 nm. During cloudless days, maximum UV-B radiation at 4 m depth was measured at solar noon (off Belize: 0.75 $W m^{-2}$, in the northern Adriatic Sea: 0.15 $W m^{-2}$).



exposing surface water from a highly stratified water column to surface solar radiation levels on cloudless days for 4 h. Before and immediately after exposure to natural radiation levels, thymidine incorporation into bacterial DNA was measured as well as 4 and 8 h after exposure to solar radiation (Fig. 2). After 4 h of exposure to surface light regimes, about 48% of suppression in bacterial growth is caused by UV-B and ~40% of the observed suppression is caused by photosynthetic active radiation (PAR) plus UV-A. After 4 h in dark, bacterial production in the bottle kept continuously in the dark and in the quartz bottle covered by a Mylar foil (PAR+UV-A) reached almost identical rates, indicating rapid recovery from light exposure in the absence of UV-B (Fig. 2). In the incubation flasks exposed additionally to UV-B radiation, thymidine incorporation was suppressed by about 56% even after 4 h in darkness, indicating prolonged suppression of bacterial growth after UV-B exposure.

To concentrate on possible adaptations to high UV-B radiation levels, we incubated water from the subpynocline layer (20 m depth) and compared the rate of thymidine incorporation into bacterial cells under surface water light regimes (PAR+UV-A+UV-B) with thymidine incorporation rates of surface layer bacteria (Fig. 3). After exposure for 4 h, no significant difference was detectable in thymidine incorporation rates between subpynocline waters and surface waters; we therefore conclude that surface water bacterioplankton does not exhibit any adaptations to UV-B radiation despite the fact that they are exposed daily to a certain level of UV-B.

Additionally, bacterial extracellular enzymatic activity was measured using fluorogenic substrate analogues⁶. After exposure to surface solar radiation levels for 4 h, bacterial extracellular enzymatic activity was reduced to less than one third of the value measured before exposure, whereas dissolved enzymatic activity in particle-free water was reduced to 50–60% of the values before UV-B exposure, depending on the enzymes tested (Table 1). These results indicate that enzyme expression is only repressed to a minor extent; more importantly, the enzymes are photolytically cleaved during UV-B exposure⁴. Therefore UV-B radiation suppresses bacterial activity and concurrently photolytically cleaves macromolecules and thereby substitutes

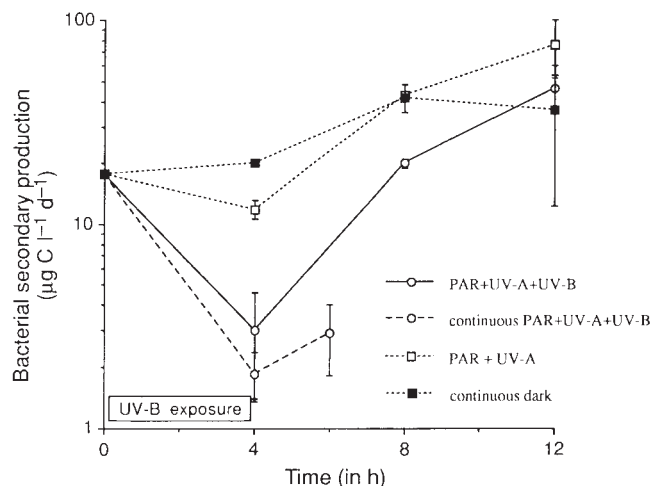


FIG. 2 Incorporation of [³H]thymidine into bacteria before and after incubation at surface solar radiation levels around noon (UV-B radiation of 0.6 W m^{-2}) for 4 h in quartz flasks (100 ml volume) (PAR+UV-A+UV-B), in borosilicate flasks covered with a Mylar foil (310 nm cutoff; PAR+UV-A). After light exposure for 4 h, all the flasks were placed in darkness; additionally, one set of flasks was either kept in continuous darkness or under solar radiation in Quartz flasks, respectively; temperature was $\sim 26^\circ\text{C}$. Each data point represents mean of two independent measurements each consisting of triplicate incubations of 5 ml subsamples with two formalin-killed blanks at *in situ* temperature in sterile Falcon tubes. Experiments were done at the Center for Marine Research, Ruder Boskovic Institute, Rovinj, Croatia.

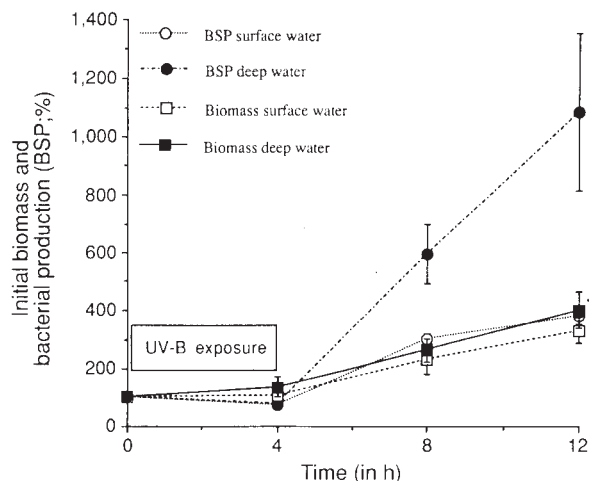


FIG. 3 Comparison of the response of bacterioplankton from subpynocline waters (from 20 m depth, temperature 12.4°C) and from surface waters (0.5 m depth, temperature 27.5°C) to surface solar radiation (incubated in quartz glass flasks around noon: 10:00 to 14:00 on a cloudless day, incubation temperature, 19°C). Higher bacterial production in deep water samples in the dark period following PAR+UV-A+UV-B exposure is probably due to the elevated temperature. Bacterial density was enumerated by epifluorescence microscopy¹² and converted into bacterial biomass using the conversion factor of 20 femtograms carbon per cell (ref. 13). Experiments were done at the Centre for Marine Research, Ruder Boskovic Institute, Rovinj, Croatia.

bacterial enzymatic activity in cleaving dissolved organic matter (DOM).

Bacterial metabolism is closely coupled to primary production⁵. Consequently, bacterial activity is highest in the upper mixed layer of the ocean where UV-B also penetrates at radiation levels high enough to influence bacterial growth. Extrapolating our measurements on UV-B penetration, we calculate that UV-B influences bacterial activity down to a depth of 10 m at around noon. Furthermore, the 1% UV-B radiation level in tropical waters is at about 30 m depth at around noon. An inhibition of bacterial activity by about 30 to 40% (depending on whether thymidine or leucine incorporation techniques have been used) is noticeable at a UV-B radiation of 0.4 W m^{-2} ; this radiation has been measured between 10:00 and 14:00 off Belize on a cloudless day down to a depth of 3.5 m. This finding has several implications in evaluating the role of bacteria in the sea. Because even surface water bacterioplankton communities from highly stratified water columns do not have adaptive capacities to protect the cells from harmful UV-B (as shown in Fig. 3), suppression of bacterial growth due to UV-B penetration in the surface layer of the ocean should be considerable. The lack of pigmentation in bacterioplankton is surprising as phytoplankton species are able to synthesize pigments to protect the cells from harmful radiation^{1,7}.

The increase in global UV-B radiation due to ozone depletion in the stratosphere might lead not only to enhanced photolysis of macromolecules in surface waters⁴ but also to a suppression of the activity of the principal consumers of this photolytically cleaved DOM, and ultimately to an increase in potentially utilizable DOM in the surface layer of the ocean. The pathway of this DOM modified by increased UV-B radiation, however, remains speculative. □

Received 11 September; accepted 22 December 1992.

1. Worrest, R. C. & Häder, D. *Environ. Conserv.* **16**, 261–263 (1989).
2. Karentz, D. & Lütze, L. H. *Limnol. Oceanogr.* **35**, 549–561 (1990).
3. Kieber, D. J., McDaniel, J. & Mopper, K. *Nature* **341**, 637–639 (1989).
4. Mopper, K. *et al. Nature* **353**, 60–62 (1991).
5. Azam, F. & Cho, B. C. *Ecology of Microbial Communities* 261–281 (Cambridge Univ. Press, Cambridge, 1987).
6. Hoppe, H.-G. *Mar. Ecol. Prog. Ser.* **11**, 299–308 (1983).

7. Marchant, H. J., Davidson, A. T. & Kelly, G. J. *Mar. Biol.* **109**, 391–395 (1991).
8. Karner, M. & Herndl, G. J. *Mar. Biol.* **113**, 341–347 (1992).
9. Fuhrman, J. A. & Azam, F. *Mar. Biol.* **66**, 109–120 (1982).
10. Robarts, R. D. & Wicks, R. J. *Limnol. Oceanogr.* **34**, 213–222 (1989).
11. Simon, M. & Azam, F. *Mar. Ecol. Prog. Ser.* **51**, 201–213 (1989).
12. Hobbie, J. E., Daley, R. J. & Jasper, S. *Appl. Environ. Microbiol.* **33**, 1225–1228 (1977).
13. Lee, S. & Fuhrman, J. A. *Appl. Environ. Microbiol.* **53**, 1298–1303 (1987).

ACKNOWLEDGEMENTS. We thank J. A. Ott for valuable comments, M. Karner, E. Kaltenböck and H. Rath for assistance during field work, P. Peduzzi for assistance in the laboratory, and the Laboratorio di Biologia Marina di Aurisina-Trieste (Italy), the Institute 'Ruder Boskovic' Center for Marine Research at Rovinj (Croatia), and the Carrie Bow Cay Station (Belize) of the Smithsonian Institution, American Museum of Natural History, for laboratory space and hospitality. Funding support for this research was provided by the Austrian Fonds zur Förderung der wissenschaftlichen Forschung to G.J.H.

Preferential representation of the fovea in the primary visual cortex

P. Azzopardi & A. Cowey

Department of Experimental Psychology, University of Oxford,
South Parks Road, Oxford OX1 3UD, UK

THE retinal fovea, which corresponds to the central degree or so of vision, is spatially over-represented in the visual cortex. It is about 0.01% of retina area, but at least 8% of the striate cortex^{1–3}. Does this simply reflect an equivalently uneven distribution of ganglion cells in the retina^{4–7}, or is the cortical representation of the fovea preferentially expanded^{8–13}? The answer hinges on the resolution of long-standing discrepancies between the retinal and cortical magnification factors. We approached the problem in a different way, using a retrograde transneuronal tracer from cortex to retina to relate directly the number of ganglion cells projecting to marked areas of striate cortex. We report here that ganglion cells near the fovea were allocated 3.3 to 5.9 times more cortical tissue than more peripheral ones, and conclude that the cortical representation of the most central retina is much greater than expected from the density of its ganglion cells.

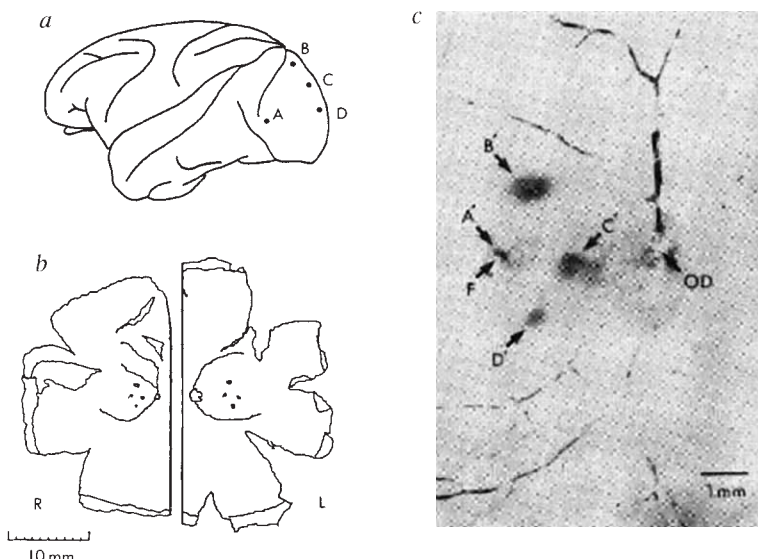
We labelled retinal ganglion cells retrogradely and trans-neuronally from the striate cortex with wheat germ agglutinin-horseradish peroxidase (WGA-HRP)¹⁴. Four microinjections of WGA-HRP were made in the left occipital operculum of two rhesus monkeys (*Macaca mulatta*) to mark and surround the foveal and perifoveal representation, and the sectioned brains and flat-mounted retinae were processed to reveal the injection sites and the loci of labelled retinal ganglion cells that projected topographically to those sites (Fig. 1). This enabled us to calcu-

TABLE 1 Area of layer IVc in the primary visual cortex allocated to P α and P β ganglion cell projections from perifoveal and peripheral retina in two rhesus monkeys

Rhesus R54				
	Number of $P\alpha + P\beta$ ganglion cells	Cortical area (mm ²)	Cortical area per ganglion cell (mm ²)	Cortical allocation ratio
Perifoveal	142,800	261.8	1.83×10^{-3}	3.29
Peripheral	1,335,800	744.4	5.57×10^{-4}	
Total	1,468,600	1,006.2		
Rhesus R151				
Perifoveal	120,100	333.3	2.78×10^{-3}	5.93
Peripheral	1,343,800	630.0	4.69×10^{-4}	
Total	1,463,900	963.3		

late the ratio of cortex to ganglion cells for the region enclosed within the labelled sites and to compare it with the same ratio obtained for the rest of the retina. The number of ganglion cells was estimated by representing their distribution as mathematical functions and integrating numerically over the polygonal region demarcated by the four labelled sites, and over the remaining area of the retina. The functions were derived from standard maps of our own previous ganglion cell counts in retinae of *M. mulatta*^{11,15} (Fig. 2). Our estimates take into account the distribution of P γ ganglion cells, which do not contribute to the retino-cortical projection. Ganglion cells representing the fovea itself are laterally displaced in the retina, but the displacement is negligible beyond 5 degrees¹⁶. Because the eccentric labelled spots in the retinae were situated at least 6.5 degrees away from the fovea, all ganglion cells projecting to the enclosed cortical region were included. The polygonal region of cortical layer IVc demarcated by the WGA-HRP injections was estimated from two-dimensional computer reconstructions of digitized serial sections. The enclosed cortical areas represented 27.8% (R54) and 34.1% (R151) of the entire area of layer IVc of the visual cortex. We confirmed these estimates by comparing the demarcated area of the lateral surface of the striate cortex encompassed by the injections, calculated from precise measurements of the distances between injection sites made during surgery, with the total area of visual cortex in the rhesus macaque². The demarcated areas represented 26% (R54) and 34.6% (R151) of the total surface area of the visual cortex, which

FIG. 1 Retrograde transneuronal labelling of retinal ganglion cells from the striate cortex. *a*, Drawing of the left cortical hemisphere showing the locations of four 0.5- μ l injections of 12–15% WGA-HRP in 0.07% kainic acid into the lateral striate cortex. Injection A was aimed at the representation of the fovea. The peripheral injections B–D were delivered 34 mm dorso-medially. *b*, Left (temporal) and right (nasal) hemiretinae of rhesus R151 showing the HRP-labelled sites. The central sites, corresponding to injection site A, were on the lip of the foveal pit. The sites corresponding to injections B–D, were located between 6.5 and 9.0 degrees away from the fovea, beyond the zone in which ganglion cells are laterally displaced¹⁶. *c*, The perifoveal region of the right retina of rhesus R151 showing the labelled sites (A'–D') corresponding to the injection sites in the cortex. F, Fovea; OD, optic disc.



is in close agreement with the values obtained after histological processing.

The average area of layer IVc of visual cortex receiving a projection from a single ganglion cell in any part of the retina provides an index of cortical allocation. Our measurements give values of $1.83 \times 10^{-3} \text{ mm}^2$ per ganglion cell (rhesus R54) and $2.78 \times 10^{-3} \text{ mm}^2$ per ganglion cell (rhesus R151) in the central

few degrees of retina, and $5.57 \times 10^{-4} \text{ mm}^2$ per ganglion cell (R54) and $4.69 \times 10^{-4} \text{ mm}^2$ per ganglion cell (R151) in the remaining region (Table 1). If cortical allocation exactly mirrored the ganglion cell distribution across the retina, the ratio of perifoveal to peripheral cortical allocation should be 1.0. But we find cortical allocation ratios of 3.29 (R54) and 5.93 (R151) in favour of the foveal-central retinal projection.

For these ratios to depart erroneously from 1.0 would require us to have overestimated the area of the enclosed cortex, or underestimated the number of enclosed retinal ganglion cells, by factors of 2.7 (R54) and 4.2 (R151). The likeliest sources of error would be to underestimate ganglion cell density in the central retina, or overestimate it in the peripheral retina, in the standard distribution function. We are confident of the reliability of these data for the following reasons. First, the estimated total number of ganglion cells in the standard retina (1.52×10^6), based on data obtained from flat-mounted retinæ sampled at 500 μm intervals, is similar to the number of optic axons in the optic nerve (1.50×10^6)¹⁷. Second, the peak density in the central retina (39.2×10^3 ganglion cells mm^{-2}) was similar to earlier estimates obtained from sections (40.0×10^3 ganglion cells mm^{-2})¹⁸. Third, the number of contaminating displaced amacrine cells in the ganglion cell layer, found in substantial numbers in the peripheral retina¹⁹, and accounted for in the original counts, was similar to the number of displaced amacrine cells estimated by us in a hemiretina of a monkey in which all ganglion cells had been destroyed by severing the optic tract 2 years earlier. But more recent data obtained from *M. fascicularis*^{6,7}, in which the foveal ganglion cell densities were estimated from sections and displaced amacrine cells were explicitly identified using immunocytochemistry, suggest that the peak density near the fovea might be somewhat higher, and the average density in the peripheral retina lower, than the earlier estimates of *M. mulatta*^{11,18}. Could the differences reported account for the apparent discrepancy between retinal and cortical magnification factors which, as we have just shown, are reflected by the cortical allocation ratios? We tested this by calculating cortical to ganglion cell ratios using distribution functions derived from the more recent data, and obtained cortical allocation ratios of 2.48 (R54) and 4.18 (R151). These are only slightly smaller than the values we obtained from our own data.

Preferential allocation of cortical area to more central retinal projections is consistent with the results of some previous anatomical and electrophysiological experiments^{8,11,20}. Because the striate cortex is homogeneous in terms of thickness, layering and cell density²¹⁻²⁴, it implies that more cortical volume is allocated on average to the more central retinal outputs¹¹. There is evidence that many psychophysical measurements, including measurements of hyperacuity, stereoaquity, colour, form and motion²⁵⁻²⁹, cannot be normalized across the visual field on the basis of retinal magnification factors alone, indicating cortical specialization of visual processing in the central visual field. Perhaps expansion of the foveal representation accommodates additional neural circuitry needed for the implementation of these processes. It might also be asked whether the preferential magnification of the foveal inputs that we have just demonstrated is determined ontogenetically or whether, as in the intact somatosensory system³⁰, it could be influenced by patterns of sensory experience. □

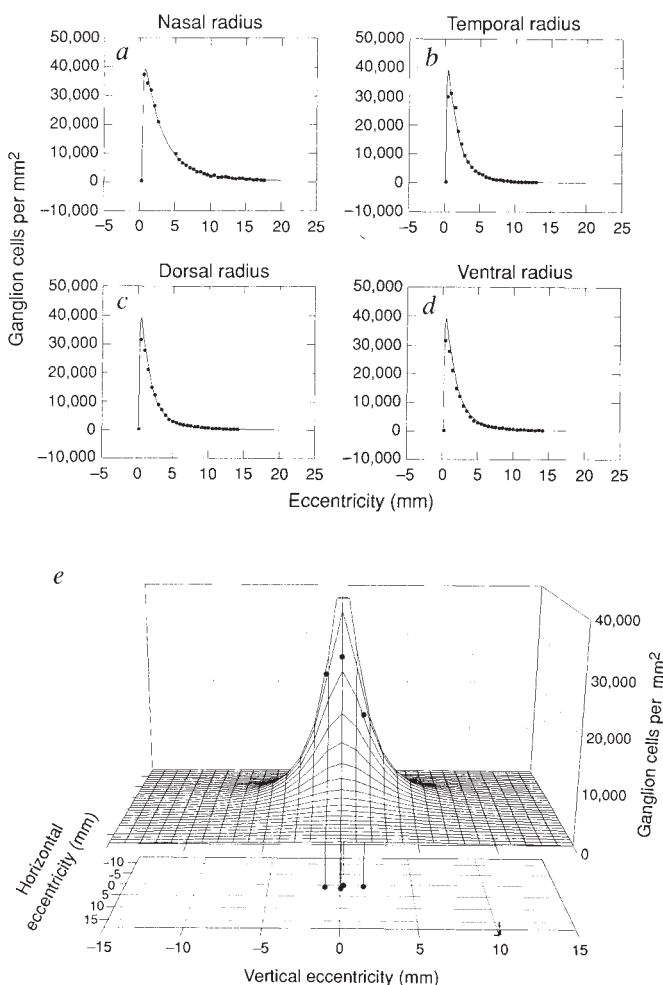


FIG. 2 Mathematical representation of the distribution of ganglion cells in the retina of *M. mulatta*, derived from standard data^{11,15}. a, Ganglion cell density as a function of eccentricity along the nasal axis. The dots represent the original data; the smooth line represents a function of the form

$$y = a_1 \cdot e^{(-l_1 \cdot x)} + a_2 \cdot e^{(-l_2 \cdot x)} + a_3 \cdot e^{(-l_3 \cdot x)} \quad (1)$$

(where x is eccentricity; $a_1 - a_3$, $l_1 - l_3$ are constants) fitted to the data using a least-squares method. b-d, Ganglion cell density as function of eccentricity along the temporal, dorsal and ventral axes. The dots represent the original data; the lines represent the same function (equation (1)) transformed by using known anisotropy to scale x . e, Ganglion cell density over the whole retina. Densities at points off the principal axes were interpolated using the elliptical function

$$r^2 = x^2/a^2 + y^2/b^2 \quad (2)$$

(x and y represent the nasotemporal and dorsoventral coordinates centred on the fovea; a and b represent the anisotropy with respect to the nasal axis), and substituting r for x in equation (1). From this function, the number of ganglion cells found in any specified patch of the retina may be evaluated numerically. In the lower part of this figure the dots indicate the two-dimensional locations of the labelled spots in the right retina of R151. These are projected onto the surface of the ganglion cell density function in the upper part of the figure. In this case only the projections of the eccentric spots are visible. The projection of the near-foveal spot, which is located inside the lip of the foveal pit, is occluded.

Received 22 June; accepted 23 December 1992.

1. Talbot, S. A. & Marshall, W. H. *Am. J. Ophthalmol.* **24**, 1255-1264 (1941).
2. Daniel, P. M. & Whitteridge, D. *J. Physiol.* **159**, 203-221 (1961).
3. Cowey, A. *J. Neurophysiol.* **27**, 366-393 (1964).
4. Drasdo, N. *Nature* **266**, 554-556 (1977).
5. Rovamo, J. & Virsu, V. *Expl Brain Res.* **37**, 495-510 (1979).
6. Wässle, H., Grünert, U., Röhrenbeck, J. & Boycott, B. B. *Nature* **341**, 643-646 (1989).
7. Wässle, H., Grünert, U., Röhrenbeck, J. & Boycott, B. B. *Vision Res.* **30**, 1897-1911 (1990).
8. Connolly, M. & Van Essen, D. C. *J. comp. Neurol.* **226**, 544-564 (1984).
9. Maipell, J. G. & Baker, F. H. *J. comp. Neurol.* **161**, 569-594 (1975).
10. Myerson, J., Manis, P. B., Mizeen, F. M. & Allman, J. M. *Science* **192**, 855-857 (1977).
11. Perry, V. H. & Cowey, A. *Vision Res.* **25**, 1795-1810 (1985).

12. Schein, S. J. *J. comp. Neurol.* **269**, 479–505 (1988).
13. Silveira, L. C. L., Picanco-Diniz, C. W., Sampaio, L. F. S. & Oswaldo-Cruz, E. *Vision Res.* **29**, 1471–1483 (1989).
14. LeVay, S. & Voigt, T. *Expl Brain Res.* **82**, 77–81 (1990).
15. Perry, V. H. & Cowey, A. *Neuroscience* **12**, 1125–1137 (1984).
16. Perry, V. H. & Cowey, A. *Neuroscience* **25**, 225–236 (1988).
17. Potts, A. M. et al. *Invest. Ophthalmol.* **11**, 980–988 (1972).
18. Rolls, E. T. & Cowey, A. *Expl Brain Res.* **10**, 298–310 (1970).
19. Perry, V. H. *Neuroscience* **6**, 931–944 (1981).
20. Dow, B. M., Snyder, A. Z., Vautin, R. G. & Bauer, R. *Expl Brain Res.* **44**, 213–228 (1981).
21. Hubel, D. H. & Wiesel, T. N. *J. comp. Neurol.* **158**, 295–305 (1974).
22. O'Kusky, J. & Collinier, M. *J. comp. Neurol.* **210**, 278–290 (1982).
23. Rockel, A. J., Hiorns, R. W. & Powell, T. P. S. *Brain* **103**, 221–244 (1980).
24. Powell, T. P. S. & Hendrickson, A. E. *Brain Res.* **216**, 409–413 (1981).
25. Levi, D. M., Klein, S. A. & Aitsebaomo, A. P. *Vision Res.* **27**, 581–597 (1985).
26. Fendick, M. & Westheimer, G. *Vision Res.* **23**, 145–150 (1983).
27. Blatherwick, P. & Hallet, P. E. in *Seeing Contour and Colour* (eds Kulikowski, J. J., Dickinson, C. M. & Murray, I. J.) 274–281 (Pergamon, Oxford 1989).
28. Strasburger, H., Harvey, L. O. Jr & Rentschler, I. *Percept. Psychophys.* **49**, 495–508 (1991).
29. Levi, D. M., Klein, S. A. & Aitsebaomo, A. P. *Vision Res.* **24**, 789–900 (1984).
30. Jenkins, W. M., Merzenich, M. M., Ochs, M. T., Allard, T. & Guic-Robles, E. *J. Neurophysiol.* **63**, 82–104 (1990).

ACKNOWLEDGEMENTS. We thank C. Healy-Yorke for histological assistance. This work was supported by the MRC.

A sensory role for neuronal growth cone filopodia

Roger W. Davenport*, Ping Dou, Vincent Rehder & S. B. Kater

Program in Neuronal Growth and Development,
Department of Anatomy and Neurobiology, Colorado State University,
Fort Collins, Colorado 80523, USA

THE dynamic nature of neuronal growth cone filopodia led to the suggestion that the primary function of filopodia is to sample their immediate environment^{1–3}, responding to and transducing environmental signals that affect growth cone behaviour^{4–6} and shape^{7–11}. Filopodia seem well suited to serve as antenna-like sensors, their broad span allows sampling of information over a greatly enhanced radius, and forward-projecting filopodia encounter potential cues in the molecular terrain long before the advancing growth cone itself¹². Filopodia in culture can serve structural roles¹³, exert mechanical tension^{14–16} and selectively adhere to their surrounding^{17–20}. Whether or not filopodia have a general sensory role has not been tested directly, largely because of their small size, which limits an electrophysiological approach, and their integral relationship with the parent growth cone, which prevents resolution of their different functions. Here we use surgical procedures to isolate individual filopodia from their parent growth cone and, by monitoring their morphology and calcium second messenger systems, we show that neuronal growth cone filopodia contain signal transduction mechanisms that allow autonomous responses and the transmission of distant environmental information to their parent growth cone.

Previous experiments have primarily examined the response of growth cones with their full complement of filopodia, and have failed to resolve the responsiveness of the filopodia versus the growth cone proper. We have directly tested the sensory role of filopodia using identified neurons from the pond snail *Helisoma*. Neurons were injected with the calcium indicator Fura-2 and filopodia subsequently isolated from their parent growth cones by careful transection with a glass micropipette (Fig. 1). Isolated filopodial membranes resealed, as indicated by their continued fluorescence. Moreover, isolated filopodia maintained their elongated form (Fig. 2a), and when stained with rhodamine-phalloidin revealed no differences in actin distribution when compared with attached filopodia. Furthermore, intracellular calcium concentrations ($[Ca^{2+}]_i$) in isolated filopodia were indistinguishable from those within the growth

cone proper (81 ± 6 nM, mean $\pm$ s.e.m., $n = 92$ filopodia; 93 ± 13 nM, $n = 25$ growth cones; $P > 0.6$). We tested the responsiveness of these isolated filopodia to stimuli known to alter growth cone $[Ca^{2+}]_i$ and behaviour.

In several types of neuronal growth cones from different species, agents causing depolarization can increase $[Ca^{2+}]_i$ within the body of growth cones by increasing calcium influx^{12,21–24}; for example, direct depolarization by an electric field can cause complete growth cone behavioural changes as a result of increased $[Ca^{2+}]_i$ ^{24,25}. When we applied depolarizing fields to isolated filopodia, there were significant rises in $[Ca^{2+}]_i$; when the stimulus terminated, $[Ca^{2+}]_i$ in filopodia returned to rest values (Fig. 3a). Given that there are voltage-dependent channels on both the soma and growth cones of these neurons²⁶, and that response to an electric field in other systems can be explained by voltage-dependent channels on growth cones²⁴, our results indicate that these channels are probably present on filopodia as well. The fact that $[Ca^{2+}]_i$ is restored after the stimulus is terminated also suggests that there are clearance and/or buffering mechanisms in filopodia as part of a homeostatic machinery to maintain filopodia at basal rest levels. Thus, filopodia can independently respond to changes in membrane polarization, a potentially important prerequisite for coupling external stimuli to changes in internal messengers.

Excitatory neurotransmitters are another class of stimuli that depolarize, raise $[Ca^{2+}]_i$ and inhibit growth cone motility^{27,28}. Addition of dopamine or serotonin, for example, raises *Helisoma* growth cone $[Ca^{2+}]_i$ and causes an accompanying inhibition of outgrowth^{22,27}. To determine whether individual filopodia can

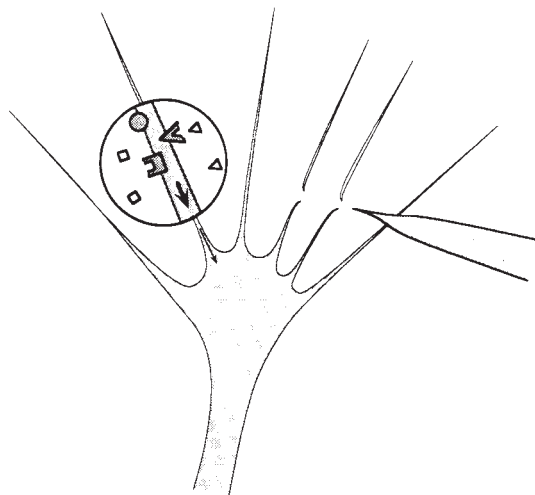


FIG. 1 Schematic model of a growth cone representing its complement of filopodia as environmental sensors. The isolation technique used to test this model is shown. Neuronal growth cones contain receptive systems that respond to a wide variety of guidance signals in their environment and affect growth cone motility through second messenger pathways. Previously, the location of these receptive systems, whether on filopodia or on the body of the growth cone or on both, was not clear. If filopodia contain receptive systems and second messenger cascades (inset), they could extend the ability of neuronal growth cones to read environmental cues by several fold and provide a highly discrete view of the surrounding environment. The model suggests that directional guidance information is conveyed directly by extracellular guidance cues to filopodia, which may integrate this information with other stimuli and transmit the information to the growth cone proper. Filopodia were mechanically isolated from the growth cone with a glass micropipette (shown for two filopodia). Filopodia isolated at differing distances from their parent growth cones were similarly responsive to applied stimuli. This result indicates that at least the distal tips of filopodia, and possibly the entire length of filopodial surfaces, were responsive to stimuli. The response of isolated filopodia, which maintained morphological integrity and a relatively stable and low $[Ca^{2+}]_i$ of less than 150 nM, to factors that affect growth cone $[Ca^{2+}]_i$ and motility was then tested.

* Present address: Max-Planck Institut für Entwicklungsbiologie, Spemannstr. 35/11, 7400 Tübingen, Germany.

respond to such stimuli, isolated filopodia were exposed to neurotransmitters: dopamine gave a significant rise in $[Ca^{2+}]_i$ (from 77 ± 10 nM, to 318 ± 43 nM; $n = 36$; $P < 0.005$), and serotonin also caused a large rise in $[Ca^{2+}]_i$ (Figs 2b, 3b). This effect was entirely dependent upon the presence of extracellular calcium: in calcium-free medium (nominally zero calcium, replaced by magnesium, plus 1 mM EGTA), $[Ca^{2+}]_i$ did not significantly change upon the addition of serotonin ($P > 0.6$, $n = 14$). Thus, filopodia can respond independently with an influx of calcium to physiological agents such as neurotransmitters.

Growth cones can integrate information from more than one signal at a time^{29,30}. For example, the simultaneous presence of the hyperpolarizing neurotransmitter, acetylcholine, can negate the rise in $[Ca^{2+}]_i$ that is the response to serotonin in these growth cones²⁹. Figures 2b and 3b show that in isolated filopodia the large rise in $[Ca^{2+}]_i$ in response to serotonin application can also be negated when acetylcholine is added. Pretreatment with

acetylcholine significantly diminishes the effect of serotonin; $[Ca^{2+}]_i$ rises to only 26% of that measured after addition of serotonin alone ($n = 13$; $P < 0.005$). Thus, individual filopodia can respond to multiple stimuli in a manner that allows direct signalling to the rest of the growth cone.

An important property of filopodia is their high surface area-to-volume ratio, which could confer high sensitivity. Assuming a filopodial volume of 10^{-15} l, only 60 calcium ions will raise the $[Ca^{2+}]_i$ by 100 nM. To compare the rise in $[Ca^{2+}]_i$ in isolated filopodia to that in attached filopodia and the growth cone proper, we used the calcium ionophore 4-bromo-A23187 to uniformly increase calcium influx³¹. The resultant rise in $[Ca^{2+}]_i$ was indeed significantly greater in isolated filopodia than in their parent growth cones (Fig. 3c), as predicted from their relatively large surface area-to-volume ratio. These results demonstrate what is potentially the most important advantage of relegating a sensory function to the filopodia, namely, the increased sensitivity afforded by the purely physical dimensions of these structures.

Excitatory neurotransmitters decrease the length and number of filopodia on intact growth cones²⁷. Using isolated filopodia we investigated whether this morphological response was an intrinsic property of filopodia, which could then have an independent contractile function as well as a sensory role. Application of dopamine significantly decreased the length of isolated filopodia in a dose-dependent manner (Fig. 4), as did serotonin (Fig. 2b). These results demonstrate that filopodia can act as autonomous units as well as convey information to the growth cone itself.

Filopodia are therefore both distant receptive extensions of growth cones and autonomous responsive units. They can act as efficient sensory probes because they are small, motile and transient structures which may require less energy to extend than the much larger mass of the growth cone; their large surface area-to-volume ratio compared with the growth cone proper makes them highly sensitive to environmental signals; and as they are discrete entities extending in a fan-like fashion across the front of the growth cone, they enable growth cones to sample information from several different areas simultaneously and discretely and therefore to compare inputs over a broad radius. There is probably a very local system of spatial discrimination and responsiveness within neuronal growth cones: the filopodia

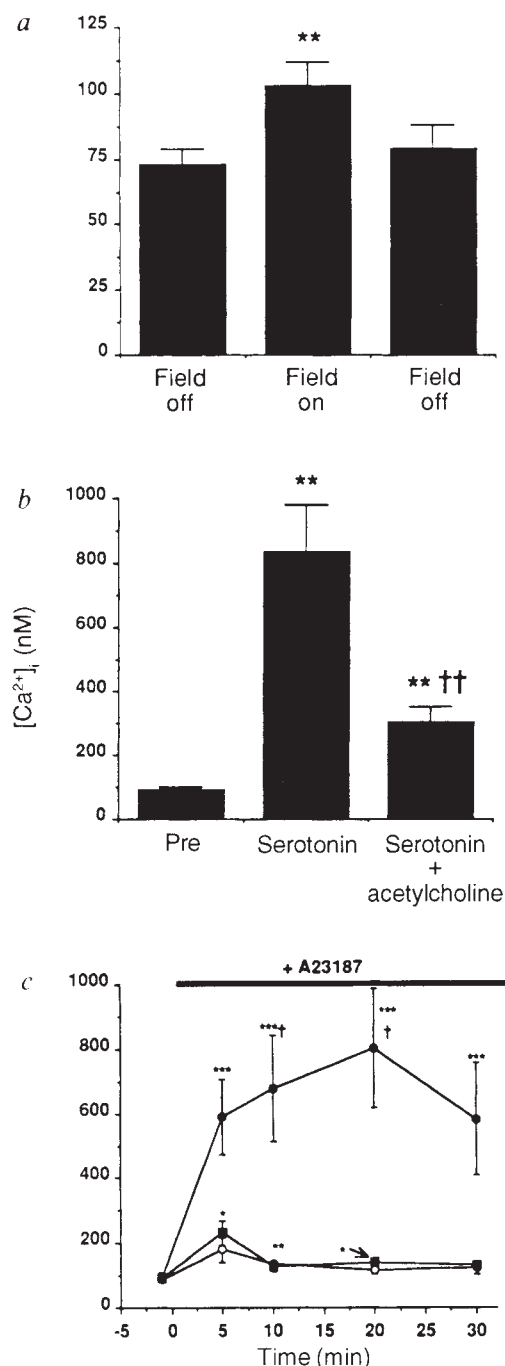


FIG. 3 Extracellular electric fields (a), the excitatory neurotransmitter serotonin (b) and the calcium ionophore 4-bromo-A23187 (c) increased $[Ca^{2+}]_i$ in isolated filopodia. **a**, Focally applied electric fields produced a rapid rise from rest $[Ca^{2+}]_i$ (73 ± 6 nM, mean $\pm$ s.e.m.; $n = 21$) within isolated filopodia. The average peak $[Ca^{2+}]_i$ reached during the first 10 s of field application was 103 ± 9 nM ($P < 0.005$). Electric field stimulation was terminated after 20–60 s. Within minutes, $[Ca^{2+}]_i$ began decreasing and was not significantly different from rest levels by 10 min (79 ± 9 nM; $P > 0.5$). **b**, Application of serotonin elicited a rise in $[Ca^{2+}]_i$ in filopodia within minutes (from 93 ± 9 nM to 837 ± 142 nM; $n = 34$; $P < 0.005$). $[Ca^{2+}]_i$ remained significantly elevated as long as serotonin was present (over 30 min, $P < 0.01$ at all time points). Addition of acetylcholine in the presence of serotonin reduced $[Ca^{2+}]_i$ (306 ± 46 nM, $n = 14$; $P < 0.005$). $[Ca^{2+}]_i$ remained significantly higher than rest levels ($P < 0.005$). **c**, Addition of 4-bromo-A23187 (0.5μ M) produced a more pronounced rise in $[Ca^{2+}]_i$ in isolated filopodia (open circles) versus attached filopodia (filled circles) or their parent growth cone proper (squares). No difference was observed between the $[Ca^{2+}]_i$ in the growth cone proper and their attached filopodia at any time point ($P < 0.08$ at 20 min, $n = 10$; $P < 0.6$ at all other time points). This is probably due to the continuity of two structures, with the growth cone acting as a sink for changes in $[Ca^{2+}]_i$. Measurements were made from 38 isolated filopodia, 11 attached filopodia and 12 growth cones. Significance levels: * $P < 0.05$; ** $P < 0.005$; *** $P < 0.002$, compared to pretreatment; † $P < 0.05$ isolated versus attached filopodia; †† $P < 0.005$ $[Ca^{2+}]_i$ in serotonin plus acetylcholine compared to $[Ca^{2+}]_i$ in the presence of serotonin alone; student's *t*-test. Arrow indicates significance applies to square, not open circle.

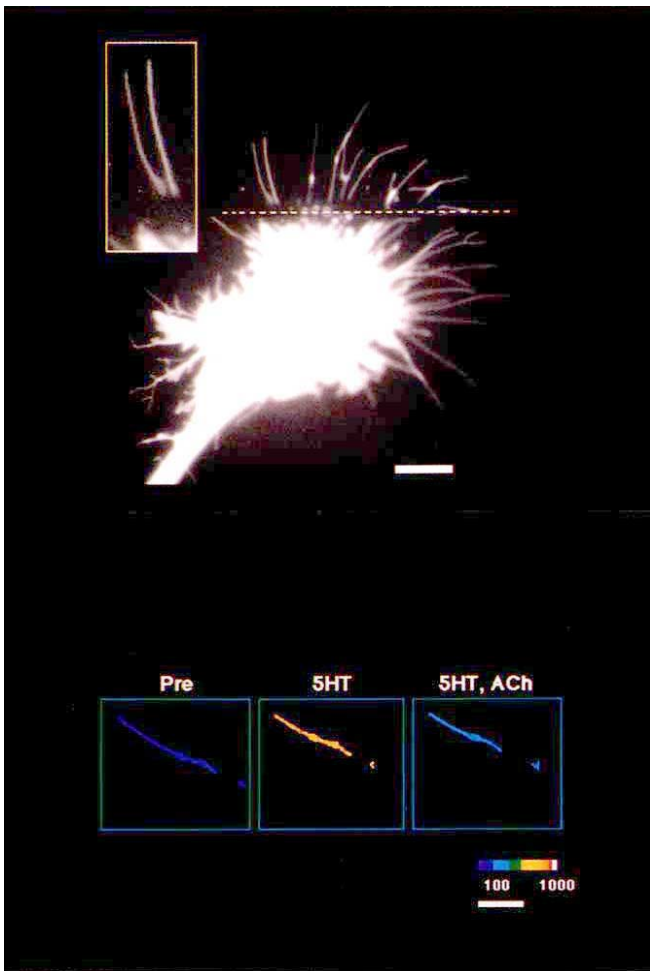


FIG. 2 $[Ca^{2+}]_i$ in filopodia was monitored to assess the responsiveness of filopodia to stimuli that affect growth cones. *a*, Emission of a Fura-2-loaded growth cone from 380 nm excitation demonstrates the ability to detect signals from isolated filopodia (enlarged in inset). *b*, Application of serotonin (5HT) increases $[Ca^{2+}]_i$ in isolated filopodia; subsequent addition of acetylcholine (ACh) decreases $[Ca^{2+}]_i$. Neurotransmitters were gently applied circumferentially to a 35-mm Falcon culture dish containing ~2 ml medium (final concentration, 50 μ M). Culture medium added alone in this way had no effect on $[Ca^{2+}]_i$. Images were obtained about 10 min after addition of neurotransmitters. Scale bars, 10 μ m.

METHODS. Filopodia were isolated from identified *Helisoma* neurons B19, plated 1–2 days before experimentation according to ref. 32. Somata were loaded with the calcium indicator Fura-2 as described^{12,25}. About 30 min later, filopodia were transected and subsequently imaged with a 100 $\times$ objective (Nikon; Fluor 1.30 oil) via a chilled CCD camera (Photometrics CH250) and digitized with software written within the laboratory to a Macintosh II fx (Datatranslation, Quick Capture frame grabber board). Filopodia were exposed for 0.5–1.2 s at two excitation wavelengths (350 ± 5 nm and 380 ± 5 nm interference filters); fluorescent emission was filtered with a 495 nm longpass filter. Experiments were done only on filopodia exceeding a stringent signal-to-noise ratio criterion: at each excitation wavelength, background fluorescence was subtracted from the filopodial signal and this quantity divided by twice the standard deviation of the background fluorescence. Only filopodia with signal-to-noise ratios greater than 3 were used for experimentation; no differences were observed among filopodia with signal-to-noise ratios ranging between 3 and 15. Calcium concentrations were estimated as described^{33,34}. Values for the imaging system were: R_{min} 0.45, R_{max} 2.92, $Kd\beta$ 1.046.

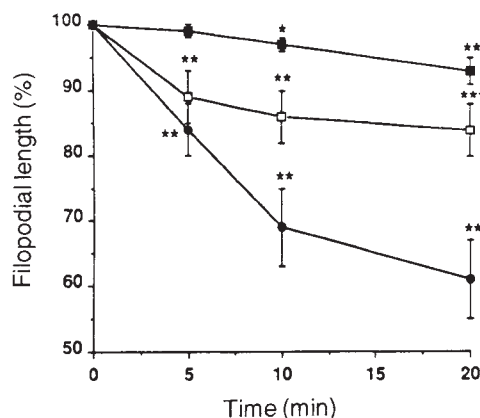


FIG. 4 Autonomous effector, as well as sensory, functions for filopodia. Stimuli that raise $[Ca^{2+}]_i$ evoke shortening of isolated filopodia. Upon application of the excitatory neurotransmitter dopamine, the length of isolated filopodia decreased relative to their respective pretreatment lengths. Measurements were made from phase-contrast images stored digitally on a Macintosh fx. Dopamine concentrations were 5 μ M (filled circles; $n=15$), 10 μ M (open squares; $n=24$), and 50 μ M (filled circles; $n=15$), respectively. At 20 min the amount of shortening with 50 μ M dopamine was significantly different from values at 5 μ M and 10 μ M ($P < 0.001$, $P < 0.05$, respectively). Significance levels: * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$, compared to pretreatment.

contain the machinery to sense and respond to the environment and convey information to the growth cone proper, which then acts as an integrator serving the full span of the filopodia and as the site of organization of motility for the elongating neurite. Thus, filopodia are a fundamental unit of organization for neuronal path-finding. □

Received 14 September; accepted 14 December 1992.

1. Sperry, R. W. *Proc. natn. Acad. Sci. U.S.A.* **50**, 703–710 (1963).
2. Bonhoeffer, F. & Gierer, A. *Trends Neurosci.* **7**, 378–381 (1984).
3. Tessier-Lavigne, M. & Placzek, M. *Trends Neurosci.* **14**, 303–310 (1991).
4. Letourneau, P. C., Kater, S. B. & Macagno, E. R. *The Nerve Growth Cone* 1–535 (Raven, New York, 1991).
5. Bandtlow, C. T. Z. & Schwab, M. E. *J. Neurosci.* **10**, 3837–3848 (1990).

6. O'Connor, T. P., Duerr, J. S. & Bentley, D. *J. Neurosci.* **10**, 3935–3946 (1990).
7. Bovolenta, P. & Mason, C. *J. Neurosci.* **7**, 1447–1460 (1987).
8. Taghert, P. H., Bastiani, M. J., Ho, R. K. & Goodman, C. S. *Dev. Biol.* **94**, 391–399 (1982).
9. Tosney, K. W. & Landmesser, L. T. *J. Neurosci.* **5**, 2345–2348 (1985).
10. Eisen, J. S., Myers, P. Z. & Westerfield, M. *Nature* **320**, 269–271 (1986).
11. Holt, C. E. *J. Neurosci.* **9**, 3123–3145 (1989).
12. Rehder, V. & Kater, S. B. *J. Neurosci.* **12**, 3175–3186 (1992).
13. Goldberg, D. J. & Burmeister, D. W. *J. Cell Biol.* **103**, 1921–1931 (1986).
14. Nakai, J. & Kawasaki, Y. Z. *Zellforschung* **51**, 108–122 (1959).
15. Heidemann, S. R., Lamoureux, P. & Buxbaum, R. E. *J. Cell Biol.* **111**, 1949–1957 (1990).
16. Bray, D. *Nature* **244**, 93–96 (1973).
17. Bray, D. *J. Cell Sci.* **37**, 391–410 (1979).
18. Hammarback, J. A. & Letourneau, P. C. *Dev. Biol.* **117**, 655–662 (1986).
19. Gundersen, R. W. *J. Neurosci. Res.* **21**, 298–306 (1988).
20. Letourneau, P. C. & Shattuck, T. A. *Development* **105**, 505–519 (1989).
21. Silver, R. A., Lamb, A. G. & Bolsover, S. R. *Nature* **343**, 751–754 (1990).
22. Cohan, C. S., Connor, J. A. & Kater, S. B. *J. Neurosci.* **7**, 3588–3599 (1987).
23. Lipscombe, D. et al. *Proc. natn. Acad. Sci. U.S.A.* **85**, 2398–2402 (1988).
24. Bedlack, R. S. Jr, Wei, M.-D. & Loew, L. M. *Neuron* **9**, 393–403 (1992).

25. Davenport, R. W. & Kater, S. B. *Neuron* **9**, 405-416 (1992).
26. Haydon, P. G. & Man-Son-Hing, H. *Neuron* **1**, 919-927 (1988).
27. Haydon, P. G., McCobb, D. P. & Kater, S. B. *Science* **226**, 561-564 (1984).
28. Mattson, M. P., Dou, P. & Kater, S. B. *J. Neurosci.* **8**, 2087-2100 (1988).
29. McCobb, D. P., Cohan, C. S., Connor, J. A. & Kater, S. B. *Neuron* **1**, 377-385 (1988).
30. Mattson, M. P. & Kater, S. B. *Brain Res.* **478**, 337-348 (1989).
31. Mills, L. R. & Kater, S. B. *Neuron* **4**, 149-163 (1990).
32. Wong, R. G., Hadley, R. D., Kater, S. B. & Hauser, G. C. *J. Neurosci.* **1**, 1008-1021 (1981).
33. Grynkiewicz, G. T., Poenie, M. & Tsien, R. *J. biol. Chem.* **260**, 3440-3450 (1985).
34. Uto, A., Arai, H. & Ogawa, Y. *Cell Calcium* **12**, 29-37 (1991).

ACKNOWLEDGEMENTS. We thank S. Sternberg and L. Detweiler for optical imaging hardware and software developments and for discussion, and D. Giddings for help in preparing the manuscript. This work was supported by a grant from the NIH.

Retinal age pigments generated by self-assembling lysosomotropic detergents

Graig E. Eldred & Michael R. Lasky

Mason Institute of Ophthalmology, University of Missouri-Columbia, School of Medicine, Columbia, Missouri 65212, USA

A UNIVERSAL biomarker of cellular ageing in eukaryotic post-mitotic cells is the appearance over time of autofluorescent lysosomal residual bodies called age pigments or lipofuscin granules¹. Their role in the process of cellular ageing has been debated without resolution². Neither the identity nor mechanism of formation of the fluorophores has been definitively determined. A post-mitotic cell type that accumulates large quantities of age pigments is the ocular retinal pigment epithelium². We have now identified the major orange-emitting fluorophore of these pigments using fast-atom bombardment tandem mass spectrometry with collisional activation analysis⁴. It is an amphoteric quaternary amine that arises as a Schiff base reaction product of retinaldehyde and ethanolamine. This compound should display lysosomotropic detergent behaviour which would help explain many of the age-related changes shown in this cell. These results suggest a new role for Schiff base reaction products as lysosomotropic amines in the genesis of cellular age pigments.

The orange-emitting fluorophore of human ocular retinal pigment epithelium (RPE) was extracted and purified as described⁵. Molecular ion analysis by high-resolution fast-atom bombardment mass spectrometry (FAB MS) revealed the exact mass to

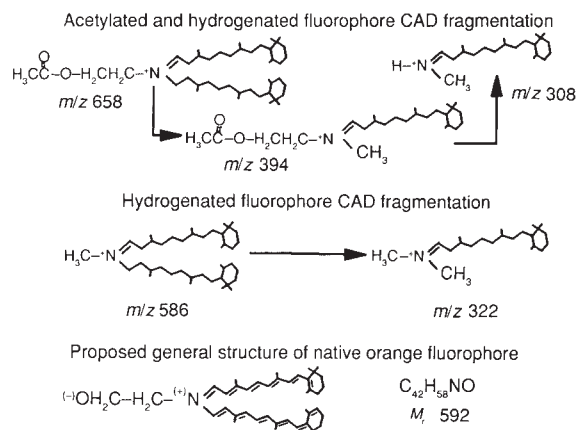


FIG. 2 Prominent positive ion, remote charge site fragments of the derivatized human orange-emitting age pigment fluorophore as revealed in the FAB CAD MS/MS spectra of Fig. 1, and its proposed general structure.

be 592.4520, which corresponds to an elemental composition of $C_{42}H_{58}NO$. An even number of hydrogens and an odd number of nitrogens implies that the nitrogen must be quaternary. The double bond equivalent (d.b.e.; number of double bonds plus rings) is 14.5; 0.5 d.b.e. is attributable to the quaternary nitrogen. Hydrogenation results in the loss of CH_2O and the addition of 24 hydrogens (that is, 12 double bonds; m/z 586 = 592 - 30 + 24). Upon acetylation, one acetyl group was added, implying that one hydroxyl group is present (m/z 634 = 592 + 42). Hydrogenation of this acetylated derivative yielded a molecule of 658 AMU (634 + 24), again confirming the presence of 12 double bonds.

The FAB fragmentation pattern is structurally ambiguous. However, FAB tandem MS with collisionally activated decomposition analysis (FAB-CAD MS/MS) of both the hydrogenated native fluorophore and the hydrogenated acetylated derivative yields a remote charge-site fragmentation pattern that is structurally unambiguous and typical of cationic surfactants⁶ (Fig. 1). The fragment envelope occurring between 658 and 394 m/z in the acetylated derivative spectrum is readily identified as the ring and branched chain of hydrogenated vitamin A. This pattern is repeated in the peaks occurring between 308 and 100 m/z . Peaks occurring between 394 and 308 m/z , and between 116 and 44 m/z are attributable to an acetylated

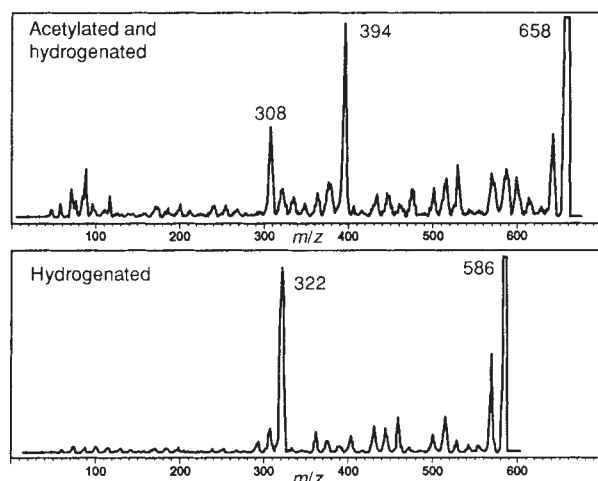


FIG. 1 Positive ion, fast-atom bombardment/collisionally activated decomposition mass spectra of the derivatized orange-emitting fluorophore of human retinal pigment epithelial age pigments. To obtain meaningful fragmentation patterns, the molecule must be hydrogenated.

METHODS. The fluorophore was purified from human RPE cells as described for fraction VIII (ref. 5). Additional purification steps involved silica gel thin-layer chromatography for 20 min with 7:1, dichloromethane:methanol and final chromatography for 30 min with 3:2:1, cyclohexane:chloroform:methanol. Purified samples were pooled and stored in methylene chloride under argon at -70°C . Working stock solutions of fluorophore contained amounts derived from at least 250 eyes of human donors of 40 years of age or older. This material ($<100\ \mu\text{g}$) was dissolved in 0.5 ml methylene chloride. For acetylation, 50 μl fluorophore stock solution was dried under nitrogen, dissolved in a mixture of acetic anhydride (50 μl) and pyridine (50 μl). The mixture was heated (50°C) for 1 h, dried under a stream of nitrogen and redissolved in 200 μl ethyl acetate. The sample was purified subsequently using flash silica chromatography on Merck Kieselgel 60 silica gel with ethyl acetate as eluent. For hydrogenation, samples were dried under a stream of nitrogen, redissolved in ethyl acetate, and added to a suspension of palladium on alumina (5%; 1 mg) in ethyl acetate (1 ml). Hydrogen gas was bubbled through this mixture for 10 min. The mixture was then filtered through a short alumina column. Fast atom bombardment (FAB) MS and collisionally activated decomposition (CAD) MS/MS spectra were obtained with a Kratos MS 50 triple-analyser mass spectrometer as described²⁴.

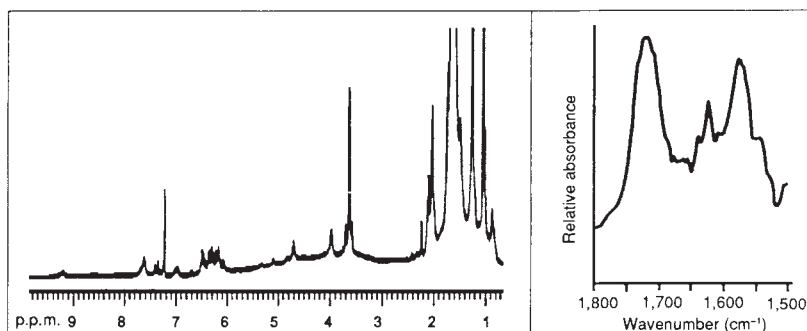


FIG. 3 Proton NMR and infrared spectra of the orange-emitting fluorophore. The imine bond is indicated at 9.2 p.p.m. in the NMR spectrum and $1,620\text{ cm}^{-1}$ in the infrared spectrum. The infrared spectrum is plotted for comparison with Schiff base spectra²⁵.

ethanolamine moiety. The structures of the main fragments are interpreted as shown in Fig. 2. The presence of an imine bond is revealed in the infrared and by proton NMR spectra (Fig. 3). The imine bond provides a bridge of conjugation in the double bond system which helps to explain its long-wavelength fluorescence emission.

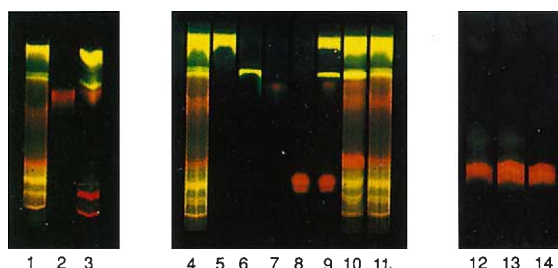
Upon arriving at this empirical structure, vitamin A aldehyde was reacted with ethanolamine *in vitro* to verify the interpretation. An orange-emitting product resulted that had identical solvent solubility, chromatographic mobility and ultraviolet-visible wavelength absorbance and fluorescence spectral characteristics to those of the extracted lipofuscin fluorophore (Fig. 4). If reaction times are longer, or the ratio of ethanolamine to retinal is increased in the synthetic reaction mixture, several other yellow-emitting compounds and a purple, non-fluorescent compound are produced which have also been observed in the extracts of human RPE age pigments. It may now be concluded that *N*-retinylidene-*N*-retinylethanolamine accounts for a substantial portion of the fluorescence of human RPE lipofuscin.

In the RPE, most of the age pigment contents are derived from phagocytosed photoreceptor outer segment membranes⁷. The source of the ethanolamine moiety in this fluorophore is very likely to be phosphatidylethanolamine⁸. Substitution of dipalmitoyl-phosphatidylethanolamine for ethanolamine in the reaction system yielded an orange-emitting fluorescent product that chromatographed to a position previously reported for the vitamin A-related, orange-emitting fluorophore which develops in the photoreceptor outer segment debris that accumulates in degenerating RCS rat outer segments⁹ (Fig. 4). From this, it appears that retinaldehyde can react with disc membrane phosphatidylethanolamine from photoreceptor outer segments to form *N*-retinylidene-*N*-retinylphosphatidylethanolamine. After the shed discs are phagocytosed by the RPE, the molecule loses its phosphoglycerolipid moiety to leave *N*-retinylidene-*N*-retinylethanolamine.

This amphoteric quaternary amine is precisely the type of compound predicted to accumulate within lysosomes¹⁰: such compounds were termed lysosomotropic amines. These are mostly nitrogenous weak bases that assume a positive charge in the acidic environment of the lysosome and become trapped. As bases they increase the intra-lysosomal pH beyond the optima of the hydrolytic enzymes. As lysosomal enzymes are inhibited, the flow of subunits to the cellular anabolic machinery is disrupted¹¹. The retinoid side chains on this molecule also give it potential surfactant properties. A variety of such lysosomotropic detergents have been described^{12,13}, including amine-based retinoid derivatives. They accumulate within the lysosome and, upon reaching a critical micelle concentration, lyse the surrounding membrane, leading to cell death¹⁴. A key characteristic of cells under the influence of subacute doses of lysosomotropic detergents is plasma membrane blebbing and the shedding of cellular cytoplasm¹⁵. Eventually, the cells vacuolate, round and die. Very similar behaviour is shown by aged RPE cells during drusen formation¹⁶, an event that marks the first clinical warning signs of pending age-related macular degeneration¹⁷.

Because age pigments appear very gradually over time¹⁸, the formation of these molecules must be a relatively rare event. *In vivo*, when *trans*-retinaldehyde is released from rhodopsin upon capturing a photon, it is rapidly converted to *trans*-retinol by the photoreceptor outer segment *trans*-specific retinol oxidoreductase¹⁹. Any retinaldehyde escaping reduction by this mechanism is likely to interact with the hydrophobic domain of the intact phospholipid membrane rather than to react with the hydrophilic ethanolamine²⁰. The Schiff base reactions would be rare, unless favoured by conditions in which the membranes are being disrupted, as is the case in the degenerating RCS rat retina, or in zinc deficiency. Zinc is a cofactor to the retinol oxidoreductases²¹. Disturbance of this enzyme activity could swing the balance of retinol/retinaldehyde toward retinaldehyde excess, which could then favour reaction with

FIG. 4 Thin-layer chromatogram of the natural and synthetic fluorophores. Lanes 1, 4 and 11, chloroform-soluble, age-related fluorophores from human retinal pigment epithelium; lanes 2 and 7, dipalmitoyl phosphatidylethanolamine/*trans*-retinaldehyde reaction product; lane 3, fluorophores associated with photoreceptor outer segment degeneration in 66-day-old RCS (Royal College of Surgeons) rats; lane 5, commercial all-*trans* retinyl palmitate; lane 6, commercial all-*trans* retinol; lane 8, ethanolamine/*trans*-retinaldehyde reaction product; lane 9, combined retinoids; lane 10, combined retinoids added to the RPE extract; lane 12, purified human orange emitting fluorophore; lane 13, combined human orange-emitting fluorophore and ethanolamine/*trans*-retinaldehyde reaction products; lane 14, purified ethanolamine/*trans*-retinaldehyde reaction product. Note that the purified synthetic ethanolamine-retinal reaction product in lanes 8 and 9 migrate lower than the native orange-emitting fluorophore in lanes 1, 4 and 11, but they co-migrate when combined in the crude extract (lane 10). When purified, the natural and synthetic orange-emitters display identical solvent solubility, chromatographic mobility, ultraviolet/visible absorbance and fluorescence characteristics (lanes 12–14). The plate is illuminated with 366 nm light. METHODS. The procedure for synthesis was modified from previously described Schiff base reaction methods^{26,27}. To a solution ($8.8 \times 10^{-3}\text{ M}$)



of all-*trans* retinaldehyde, the appropriate amine was added to achieve a molar ratio of 2:1, retinaldehyde:amine. For catalysis²⁸, the mixture was acidified with acetic or lactic acid to a pH of 5.2–5.7 and allowed to react under argon at room temperature for 20 min or longer. The reaction mix was then dried under argon, reconstituted in CH_2Cl_2 and stored at -70°C . Thin-layer chromatography was on silica gel plates using the primary running solvents described⁵.

phosphatidylethanolamine. Such a mechanism might be a factor in the beneficial effects of dietary zinc in the aetiology of age-related macular degeneration²², and in the formation of lipofuscin-like inclusion bodies in animal models of zinc deficiency²³.

Whether this or other lysosomotropic detergent amines accumulate within the age pigments of other tissues remains to be determined. Retinaldehyde is not likely to be a factor in other tissues, but other aldehydes, possibly originating from lipid oxidation, may react to form similar lysosomotropic amines by Schiff base reactions and accumulate in lysosomes, inhibiting proteolysis in the ageing cell. □

Received 1 July; accepted 18 December 1992.

- Porta, E. A. *Arch. Gerontol. Geriatr.* **12**, 303–320 (1991).
- Sohal, R. S. & Brunk, U. T. *Lipofuscin and Ceroid Pigments* (ed. Porta, E. A.) 17–29 (Plenum, New York, 1989).
- Feeney, L. *Invest. Ophthalmol. vis. Sci.* **17**, 583–600 (1978).
- Tomer, K. B. *Mass Spectrom. Rev.* **8**, 445–482 (1989).
- Eldred, G. E. & Katz, M. L. *Expl Eye Res.* **47**, 71–86 (1988).
- Lyon, P. A., Crow, F. W., Tomer, K. B. & Gross, M. L. *Analyt. Chem.* **56**, 2278–2284 (1984).

- Feeney-Burns, L. & Eldred, G. E. *Trans. ophthalm. Soc. U.K.* **103**, 416–421 (1983).
- Fliesler, S. J. & Anderson, R. E. *Prog. Lipid Res.* **22**, 79–131 (1983).
- Katz, M. L., Eldred, G. E. & Robison, W. G. Jr *J. Mech. Ageing Dev.* **39**, 81–90 (1987).
- deDuve, C. et al. *Biochem. Pharmacol.* **23**, 2495–2531 (1974).
- Seglen, P. O. *Meth. Enzym.* **96**, 737–764 (1983).
- Firestone, R. A., Pisano, J. M. & Bonney, R. J. *J. med. Chem.* **22**, 1130–1133 (1979).
- Pisano, J. M. & Firestone, R. A. *Synth. Commun.* **11**, 375–378 (1981).
- Miller, D. K., Griffiths, E., Lenard, J. & Firestone, R. A. *J. Cell Biol.* **97**, 1841–1851 (1983).
- Martinez, C. M. et al. *Cell Biol. Int. Rep.* **14**, 255–266 (1990).
- Burns, R. P. & Feeney-Burns, L. *Trans. Am. Ophthalm. Soc.* **78**, 206–225 (1980).
- Eagle, R. C. Jr *Ophthalmology* **91**, 613–625 (1984).
- Feeney-Burns, L., Hilderbrand, E. S. & Eldridge, S. *Invest. Ophthalm. vis. Sci.* **28**, 195–200 (1984).
- Zimmerman, W. F., Lion, F., Daemen, F. J. M. & Bonting, S. L. *Expl Eye Res.* **21**, 325–332 (1975).
- Robert, S., Tancrede, P., Salesse, C. & LeBlanc, R. M. *Biochim. biophys. Acta* **730**, 217–225 (1983).
- Saari, J. C. & Bredberg, L. *Biochim. biophys. Acta* **716**, 266–272 (1982).
- Newsome, D. A. et al. *Arch. Ophthalmol.* **106**, 192–198 (1988).
- Leure-duPree, A. E. & McClain, C. J. *Invest. Ophthalm. vis. Sci.* **23**, 425–434 (1982).
- Gross, M. L. et al. *Int. J. Mass Spectrom. Ion Phys.* **42**, 243–254 (1982).
- Rhoif, C., Vocelle, D. & Sandoz, C. *Res. Chem. Intermed.* **12**, 131–139 (1989).
- Palmer, B., Jumper, B., Hagan, W., Baum, J. C. & Christensen, R. L. *J. Am. chem. Soc.* **104**, 6907–6913 (1982).
- Birge, R. R., Murray, L. P., Zidovetzki, R. & Knapp, H. M. *J. Am. chem. Soc.* **109**, 2090–2101 (1987).
- Wagner, E. C. *J. org. Chem.* **19**, 1862–1881 (1954).

ACKNOWLEDGEMENTS. We thank R. Hayes, M. Gross and R. Cerny for their services, T. Wong for NMR analysis, and G. Owens for help with tissue extraction, photography and a literature search. Supported by the NIH and by a grant-in-aid from Research to Prevent Blindness, Inc. Tissues were provided through the Eye Research Foundation and the Lions Tissue Eye Bank, Columbia, Missouri. The 500 MHz NMR spectrometer was partially funded by the NSF.

Putative X-linked adrenoleukodystrophy gene shares unexpected homology with ABC transporters

Jean Mosser*, Anne-Marie Douar†, Claude-Olivier Sarde*, Petra Kioschis‡, Robert Feil*§, Hugo Moser||, Anne-Marie Poustka‡, Jean-Louis Mandel* & Patrick Aubourg†¶

* Laboratoire de Génétique Moléculaire des Eucaryotes du CNRS, INSERM Unité 184, Institut de Chimie Biologique, Faculté de Médecine, 67085 Strasbourg, France

† INSERM Unité 342, Hôpital Saint Vincent de Paul, Faculté Cochin, Université René Descartes, 82 Avenue Denfert Rochereau, 75014 Paris, France

‡ Deutsches Krebsforschungszentrum, Institut für Virusforschung, Im Neuenheimer Feld 506, Heidelberg, Germany

§ Kennedy Krieger Institute, Baltimore, Maryland 21205, USA

ADRENOLEUKODYSTROPHY (ALD) is an X-linked disease affecting 1/20,000 males either as cerebral ALD in childhood or as adrenomyeloneuropathy (AMN) in adults¹. Childhood ALD is the more severe form, with onset of neurological symptoms between 5–12 years of age. Central nervous system demyelination progresses rapidly and death occurs within a few years. AMN is a milder form of the disease with onset at 15–30 years of age and a more progressive course. Adrenal insufficiency (Addison's disease) may remain the only clinical manifestation of ALD. The principal biochemical abnormality of ALD is the accumulation of very-long-chain fatty acids (VLCFA) because of impaired β -oxidation in peroxisomes^{1,2}. The normal oxidation of VLCFA-CoA in patients' fibroblasts^{3–5} suggested that the gene coding for the VLCFA-CoA synthetase could be a candidate gene for ALD. Here we use positional cloning to identify a gene partially deleted in 6 of 85 independent patients with ALD. In familial cases, the deletions segregated with the disease. An identical deletion was detected in two brothers presenting with different clinical ALD phenotypes. Candidate exons were identified by computer analysis of genomic sequences⁶ and used to isolate complementary DNAs by exon

connection⁷ and screening of cDNA libraries. The deduced protein sequence shows significant sequence identity to a peroxisomal membrane protein of *M_r* 70K that is involved in peroxisome biogenesis and belongs to the 'ATP-binding cassette' superfamily of transporters.

As previous attempts to purify VLCF-CoA synthetase were unsuccessful, we used a 'positional cloning' approach. The ALD locus has been mapped to Xq28 (refs 8–10), where the red/green colour pigment (R/GCP) genes reside¹¹. On the basis of the high incidence (40%) of colour vision anomalies in AMN patients¹² and our earlier results¹³, we proposed that ALD and R/GCP genes could be close together. Recently, we identified an AMN patient with blue-monochromatic colour vision who had a complex rearrangement located 5' of the red-colour pigment (RCP) gene^{14,15}, which included two deletions separated by a large (>110 kb) inversion¹⁵ (Fig. 1a and b). Only the RCP gene was found in the first deletion (88 kb). No additional deletion was detected in this region in 81 other ALD patients. We postulated that the inverted segment or the second deletion were candidate loci for the ALD gene.

Probes corresponding to three of the breakpoints (BP) of this rearrangement were isolated (BP1, BP2 and BP3)¹⁵ (Fig. 1). To estimate the size of the second deletion, we used a probe 4 kb proximal to *Fr15* (deleted in patient O; ref. 15) to obtain clones from a Xq28 cosmid library¹⁶. Three overlapping clones were obtained that spanned about 90 kb and contained a cluster of rare restriction cutting sites (*Eag*I, *Bss*HII, *Sac*II) indicating the presence of a CpG island (Fig. 1c). In parallel, we cloned an *Xba*I junction fragment (X-8) corresponding to breakpoints 4 and 2 in patient O (Fig. 1a). A restriction map of X-8 showed that a 1.8 kb *Xba*I-*Eco*RI (X2) fragment contains a 1.6-kb segment on the BP4 side and included within the *Fr15* cosmid contig described above (Fig. 1b and c). Breakpoints 3 and 4 are separated by 19.2 kb, and delimit the second deletion in patient O.

To search for conserved sequences in other mammalian species and to detect additional deletions in other ALD patients, we subcloned cosmid Qc 11H12 (Fig. 1b, c). Probes TA25, TA1, TA18 and TA13b (Fig. 1c) showed cross-hybridization to various mammalian species. More important, X2, TA4, TA18 and TA13a probes detected deletions in five other ALD patients. Probe X2 allowed the detection of a junction fragment segregating with the disease in family B (Fig. 1d). In another family (Ma), probe X2 detected an identical 22-kb *Hind*III junction fragment in two brothers with different clinical phenotypes of ALD (Fig. 1e). The size of deletions ranged from 1.6 kb (patient B) to 15.3 kb (patient R) with partial overlap (Fig. 1c). Six patients had

* To whom correspondence should be addressed.

§ Present address: Department of Molecular Embryology, Cambridge Research Station, Cambridge CB2 4AT, UK.

deletions (including patient O) in a population of 85 independent ALD patients, but no deletions were found in a panel of 82 control males. These results indicate that the region contains at least part of the ALD gene.

TA25, TA18 and TA13b detected a faint 4.2-kb transcript in several tissues. Although six human oligo(dT)-primed cDNA libraries (derived from fetal or adult muscle, fetal adrenal and brain, lung and foreskin fibroblasts) were screened with these probes, no clones corresponding to the candidate region were obtained. The sequences of these probes and of TA13a were determined and examined for putative coding regions (using a computer program based on a multiple sensor-neural network approach⁶) revealed a large (>700 bp) putative protein coding sequence within TA25 and smaller open reading frames (300–400 bp) in TA18, TA13a and TA13b. The deduced amino-acid sequences showed significant sequence identity with collinearly positioned regions of human or rat 70K peroxisomal membrane protein (PMP)^{17,18}. Nested-PCR reactions using primers from the putative exons (Fig. 2) produced two fragments (Ex13 and Ex3) of sizes (645 bp and 498 bp, respectively) compatible with those expected from homology with the 70K PMP cDNA. They

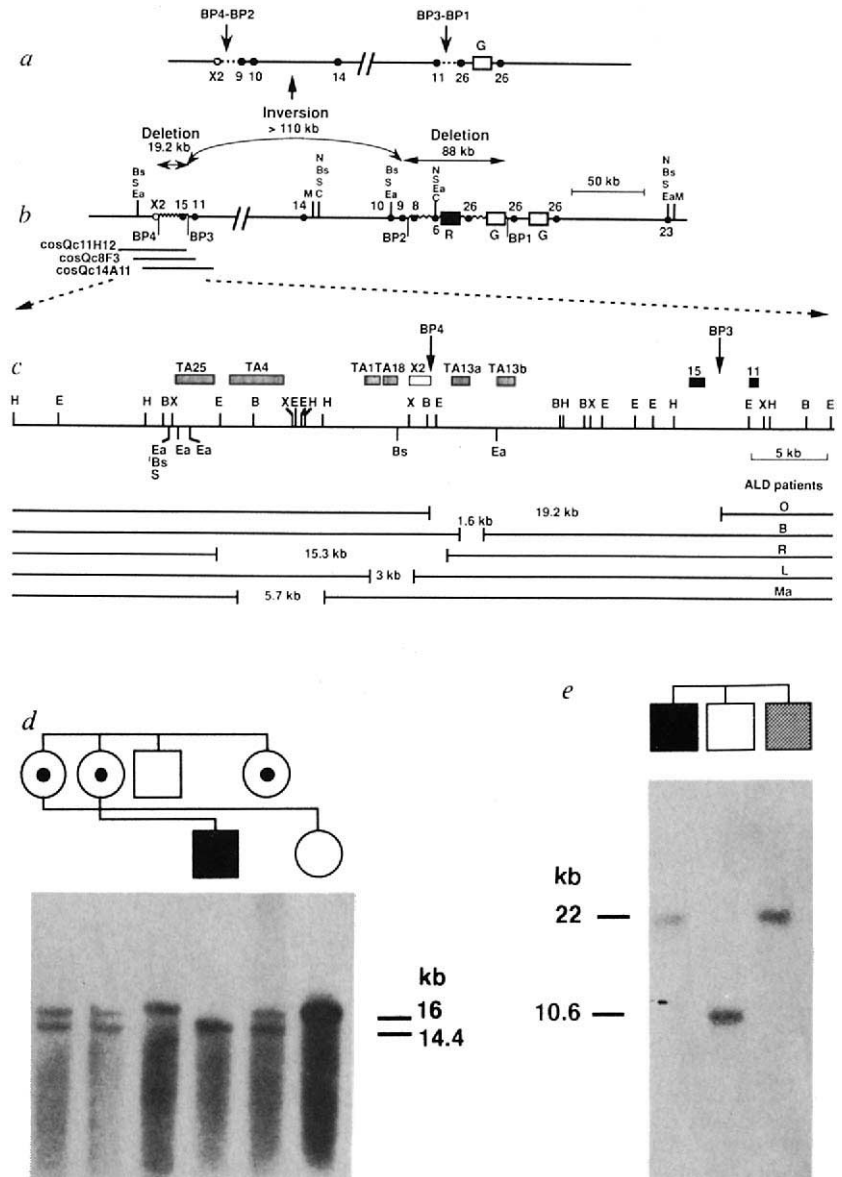
hybridized to the predicted DNA fragments in normal individuals and to the same fragments detected by genomic probes in ALD patients (Fig. 3a). Ex13 and Ex3 have been used in combination to screen a random-primed HeLa cell cDNA library¹⁹ to obtain 6 independent overlapping clones.

The 2,751-bp sequence (Fig. 2) contains the whole protein-coding sequence of 745 amino acids. The first methionine codon is preceded by an in-phase stop codon (at bp 282) and is included within a potential ribosome-binding sequence²⁰. Significant sequence identity with the 70K PMP begins at methionine 67 and ends at ~680 (corresponding to the carboxy terminus of 70K PMP). The remaining 52 amino acid residues are unique to the ALD protein (Fig. 4a).

The sequence of ALD protein could be aligned with human 70K PMP^{18,21}, with only a few deletions or insertions, and revealed a 38.5% amino-acid identity (253/659 amino acids) (Fig. 4a). When conservative amino acids substitutions are considered, the sequence similarity between the two protein sequences reached 78.9%. The two proteins show a similar hydrophobicity profiles (not shown), with a hydrophobic amino-terminal region containing potential transmembrane segments.

FIG. 1 Map of the ALD gene region and its rearrangements in patients. *a*, Chromosomal rearrangement in patient O. The joining of BP4-BP2 and BP3-BP1 were demonstrated by cloning the corresponding junction fragments (dashed lines). The extents of the two deletions (19.2 and 88 kb respectively) are indicated in *b*. Probe X2 (circle) is discussed in the text and other probes (filled circles, Fr probes) are from ref. 15. The distance between probes Fr14 and Fr11 (broken line) is not known. The green cone pigment gene (GCP) is indicated by a box (G). *b*, Map of the R/GCP genes and of the second deletion in patient O, including the rare cutter restriction sites (vertical bars, top) *EagI* (Ea), *BssHII* (Bs), *SacII* (S), *ClaI* (C), *NotI* (N) and *MluI* (M). The rare cutter sites within the pigment gene repeat unit are not marked. The extents of the two deletions (wavy lines) and the position of the 4 BPs in patient O are indicated. Probe Fr15.4 (deleted in patient O) was used to screen a Xq28-specific cosmid library, yielding 3 overlapping clones (cos Qc 11H12, cos Qc 8F3 and cos Qc 14A11). The red cone pigment (RCP) gene is indicated by the filled box (R), and the GCP gene by a box (G). *c*, Deletions detected in 5 ALD patients and restriction map of the subcloned region of ALD gene using the following enzymes: *EcoRI* (E), *HindIII* (H), *BamHI* (B), and *XbaI* (X). Rare cutter sites are indicated as in *b*. Localization of subcloned probes is shown at the top. Probe X2 (box) is a 1.8-kb *XbaI*-*EcoRI* restriction fragment derived from X-8, the second junction fragment (BP4-BP2) in patient O. TA25 (2.1 kb), TA4 (3.6 kb), TA1 (1.0 kb), TA18 (0.85 kb), TA13a (935 bp) and TA13b (252 bp) are *TaqI*-digested DNA fragments (hatched boxes) derived from subcloning of cosmid Qc 11H12. *d*, Segregation of an abnormal junction fragment detected by probe X2 in ALD family B. Probe X2 hybridizes to a 16-kb *HindIII* fragment in normal individuals (open square and circle). An abnormal 14.4-kb junction fragment was detected in an affected ALD patient (patient B; see panel *c*) and in all heterozygous females (○). *e*, Detection of the same rearranged DNA fragment in two ALD brothers with different clinical ALD phenotypes by Southern blot analysis of *HindIII*-digested DNA from 3 brothers (Family Ma; see *c*) hybridized with probe X2. An abnormal junction fragment of 22 kb is detected by X2 in a male with cerebral ALD (filled square) and his brother with Addison's disease (hatched square).

METHODS. Restriction patterns of the cosmids were analysed as described¹⁵. Cosmid Qc 11H12 was digested with *TaqI* and cloned directly in the *ClaI* site of pBluescript SK⁺ (Stratagene). X-8 was isolated from a *XbaI* genomic library constructed in bacteriophage λZAPII (Stratagene) using DNA from a somatic hybrid line containing the X chromosome of



patient O¹⁵. Gel electrophoresis, Southern blotting, probing and autoradiography were all done as described^{14,15}.

METHODS. Four probes (Ta25, Ta1, Ta18 and TA13b) derived from subcloning of cosmid Qc 11H12 were sequenced. Candidate exons, strand and frame assignments were screened by the GRAIL program (Oak Ridge, TN)⁶. Oligonucleotide primers were designed according to the coding regions that presented highest homology score to 70K PMP in TA25, TA18, and TA13b. Ex13 and Ex3 are cDNA clones obtained by exon connection⁷ in 2-step-boosted or nested polymerase reactions (30 cycles with external primers and 40 cycles with internal primers) performed on oligo(dT)-primed cDNA³⁵. Total RNA (20 µg) from a lymphoblastoid cell line was used as starting material. External primers correspond to positions 1,853–1,872 (for Ex13), and to positions 1,854–1,874 and 2,357–2,375 (for Ex3). Internal primers are indicated by arrows. Subsequent amplification products were blunt-ended by action of T4 DNA polymerase (New England Biolabs), directly cloned in pBlue-script KS⁺ vector (Stratagene), sequenced using dideoxynucleo-

Ex3

sequenced using dideoxynucleotide termination (Applied Biosystems), and analysed on an automated DNA sequencer (Applied Biosystems).

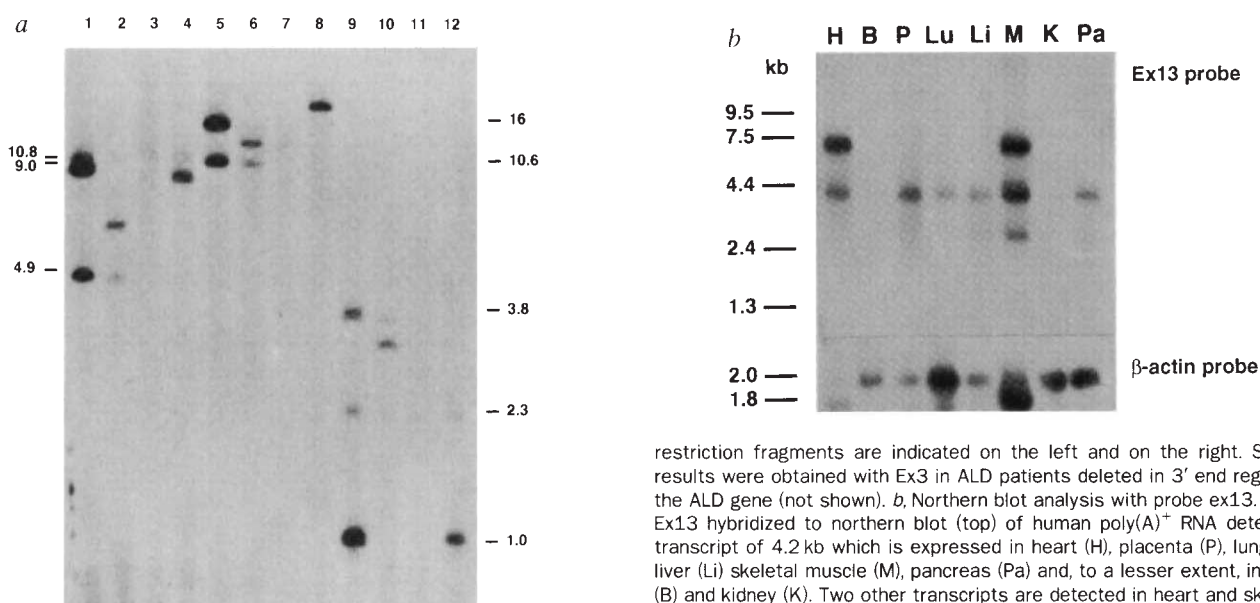


FIG. 3 a, Detection of deletions in the DNA isolated from three ALD patients using cDNA probe Ex13. DNA was digested with *Eco*RI (lanes 1–4), *Hind*III (lanes 5–8) and *Taq*I (lanes 9–12). Lanes 1, 5 and 9: normal woman; lanes 2, 6 and 10: patient L; lanes 3, 7 and 11: patient R; lanes 4, 8 and 12: patient Ma. The 10.8-kb *Eco*RI normal fragment (lanes 2–4) and the 12.5-kb *Hind*III abnormal junction fragment from patient R (lane 7) hybridize to only ~60 bp of the Ex13 probe and are thus very faint. Sizes (in kb) of normal

restriction fragments are indicated on the left and on the right. Similar results were obtained with Ex3 in ALD patients deleted in 3' end region of the ALD gene (not shown). *b*, Northern blot analysis with probe ex13. cDNA Ex13 hybridized to northern blot (top) of human poly(A)⁺ RNA detects a transcript of 4.2 kb which is expressed in heart (H), placenta (P), lung (Lu), liver (Li) skeletal muscle (M), pancreas (Pa) and, to a lesser extent, in brain (B) and kidney (K). Two other transcripts are detected in heart and skeletal muscle (6.8 kb) and in liver and skeletal muscle (2.75 kb). RNA size markers are indicated on the left. A human β -actin probe hybridized to the same northern blot (bottom panel) detects two transcripts of 2.0 and 1.8-kb, respectively.

METHODS. Gel electrophoresis, Southern blotting, probing and autoradiography were done as described⁴⁵. For *b*, human multiple-tissue northern blot was purchased from Clontech. Membranes were exposed at -70°C to X-ray film for 5 d (Ex13 or EX3 probes) or for 6 h (β -actin probe).

a

PMP70	MAAFSKYLTARNSSLAGAALFGLIHKRRRA--LGL--HGKKSCK--PPIQ--NNKEG	1-52
ALDP	MDVLSPRPARGNTLKRTAVLALAAAYGAHVYDILVRCQIAPARGLQAPAGEPTQHASGV	1-60
PMP70	KKERAVVDKVEPSRLIQILKIMVPTCKRTGTGVIIVMLVSRITYCDVMMIQNTLIES	53-112
ALDP	AAARAGMNRVFLQRLWLLRLLEPRVLCREFGTGLAHSAALVSRTEFSVYARLDORLAR	61-120
PMP70	GIIGRSKDKFRKYLNFIAAPPLISLVNPHIKYGINELKLCFVRRLTKYLYERYLQAFY	113-172
ALDP	CTARKDPRAFQWQLQLWLLIAPATHVNSAIRYLEGQALSFSSRLVAHATRLYFSQQT	121-180
PMP70	YKKGNDNRIANPDQLITQDVEKCNVVDVLSNLSKPLDVIYIFKLTSALGAQGPAA	173-231
ALDP	YRVSNMDGRIKNPDSITEDVFAAASVAHLVSNLTKPILDVAVTSYTLRAAKSRGAGT	181-240
PMP70	---SMMAYLVV--SGFLTRLRPPIGKILTEQRYKGRYRVNSRLITNSERTAFYNOK	232-286
ALDP	ANPSAIAIGVVELTANVIRAFSPKPGHIVAEARRKGLRYMISRVVANSEELAFYGGH	241-300
PMP70	REKQTVHSVFRKLVEHLNFIIRFSSMFIUSIIAKYLATVVGVLVVSRRFL-----	287-338
ALDP	VELALQSRVQDLASQNLILIRLIVYVLEQFLMKYVWSAGLIMVAVPIITATGYSFS	301-360
PMP70	--DLSPHRLKSTHSELI---RDYYSQGRMLLR--MSQALGRIVLAGREMTIRACFTARIT	339-392
ALDP	DAEAVKKAALKKKKKELVSRTEAFETARNLLTAADAIERIMSSYKVEVTELAGYTVARH	361-420
PMP70	FIMQVLKDLNKGHYERTM--VSQQRK----GIEGVQVILPLIPGAGEIITADNIIKFDHPV	393-446
ALDP	RMFOVFDVQCHFRPRPRLLEDQAQSGCTIGRSGVVRVEPIKIRGQVVDVEGILICGNT	421-480
PMP70	LATPNDVLIRDLNEFVRSGANVLICGPGCGCKSSIFRVGLGLKPLGGLRICKPERRKIF	447-506
ALDP	IVTPSGGVVASLNIIRVEGMILLITGPGCGCKSSIFRILGGLKPLTGGVLYKPPQRMF	481-540
PMP70	YVPRPQYMTLGTLDQVYVPGREGQKRRKISDILVQKRYLDNVQLHILREGQWDSVQ	507-566
ALDP	YVPRPQYMSVGSRLDQVYVPSVPMQKRYSEQDLLEATLDVVMHILHILREGQWAGD	541-600
PMP70	WMDVLSGCGKQRMAMARILYKQKQAFILDECTSAVSVDVEGYIYSHCRKVGITLPTVSR	567-626
ALDP	WMDVLSGCGKQIRGMARMYHRPKYALIDECTSAVSIDVEGIIQAQADAGIALISITIR	601-660
PMP70	KSLLKHHEYLHMDGRNVEPKITRYTVFVGS-----	627-659
ALDP	PSLKKYHHLQLQDGGGKPKKIDSAARLSITEEKQRLQQLAGIPKMRRLQELCQIL	661-720
PMP70	-----	
ALDP	GEAVA?AHVAPSPQCCGCGCAGST	721-745

b

ALDP	486-544	GENVVV--ASINIRVEEGKRLITLTPNGCGCKSSSLRVLGLWLETGGVILKPPFMMFYVYQ
HPMP70	452-510	QVTV--RDINFEVRSGRNLLTCGPNCGCKSSSLRVLGLWLETGGVILKPPFMMFYVYQ
RPMP70	452-510	QVTV--QDLSFEVRSGRNLLTCGPNCGCKSSSLRVLGLWLETGGVILKPPFMMFYVYQ
MDR1a	406-464	EMVKKLGLNLSVSGQQTLYVGNSSCGCKSSITVQVMQRIDPTECHMSVDGQGIIRITINVR
MDR1b	1049-1107	EMVKKLGLNLSVSGQQTLYVGNSSCGCKSSITVQVMQRIDPTECHMSVDGQGIIRITINVR
CFTRa	437-495	CT-PVLKDNFIRKQQLLKAAGSTGAGKYSLLMMMLLEPSECKKKSGRISFCSQFS
CFTRb	1222-1280	CGNKLLENISFSFPCQVQLLGRIGCKSSITLSLFLRLINTEGZQIDGVSWDSITLQ
PSF1	576-635	PEVVLVLCQITFTLRPEGVTLVCGPNCCKSSYAAALQNPQPTGQQLLDGKFLPQVHR
PSF2	493-552	PEVVLVLCQITFTLRPEGVTLVCGPNCCKSSYAAALQNPQPTGQQLLDGKFLPQVHR
MALK	14-73	GENVVVSKDINIDHGEFVYVFGSPSGCKSSITLSLFLRLINTEGZQIDGVSWDSITLQ
ALDP	545-596	REYMYVGLRDQ--VIYEDSVEDM-----RKGYSQDLNAILDVVHLHILHILREGGWA
HPMP70	511-562	REYMYVGLRDQ--VIYEDSVEDM-----RKGISLVLQKEYLDNVQLGHILHILREGGWA
RPMP70	511-562	REYMYVGLRDQ--VIYEDSVEDM-----RKGISLVLQKEYLDNVQLGHILHILREGGWA
MDR1a	465-521	FLREITGVSSQSPFYAT-TIAEIRYGRN--TDEIEKAKKENAYDFIMKLPHKYS
MDR1b	1108-1166	FLREITGVSSQSPFYAT-TIAEIRYGRN--TDEIEKAKKENAYDFIMKLPHKYS
CFTRa	496-538	WIMP-CYKKNITIGVSY-----DEYR-----YRSYKACQLEEDTSKFAEKD
CFTRb	1281-1336	QWRKAFVYQPKVFLSGFRKNLDPEYQMSD-----QIWKVADKVLRSVLEFPFKLD
PSF1	636-693	YLHRQVAAVGPQVYGR-SLGEIAYGLT-QPTTEITAAKMSAHSFISGLPQK
PSF2	553-610	YLHRQVAAVGPQVYGR-SLGEIAYGLT-QPTTEITAAKMSAHSFISGLPQK
MALK	74-125	ERGVCYFYSAYLYPHLSVAEEMSFGLKPAGKKKEVINERNGVGE--LQLAR--
ALDP	597-651	MCDDWLDVLSGGEKORIMARMFYHPEKALLDECTSAVS-----VEGKYSQAAKDAG
HPMP70	563-617	S-VQDWMQVLSGGEKORMMARLFYHPEKALLDECTSAVS-----VEGYYSCHCKXG
RPMP70	563-617	S-VQDWMQVLSGGEKORMMARLFYHPEKALLDECTSAVS-----VEGYYSCHCKXG
MDR1a	522-579	TLVGRGAQLSGGKORITIALVKNPKILDEEATSADTISEAVVQAA--DKARKG
MDR1b	1167-1224	TLVGRGAQLSGGKORITIALVKNPKILDEEATSADTISEAVVQAA--DKARKG
CFTRa	539-597	FVLVGGGCTLSGGKORISLARAVYDADLYLDDSPFGYDILTEKEFES-CVCKLMAN
CFTRb	1337-1394	FVLVGGGCTLSGGKORISLARAVYDADLYLDDSPFGYDILTEKEFES-CVCKLMAN
PSF1	694-753	TEVDVAGSGLSGGKQOALALALIKKPCVLLDDEATSAADNSQLQVEQLXSPERYS
PSF2	611-670	TEVDVAGSGLSGGKQOALALALIKKPCVLLDDEATSAADNSQLQVEQLXSPERYS
MALK	126-184	LLDRKPKALSGGKORVATRTLVAEPSVFLDDEPLSNKDA-RVQMRKESRLHKLRLG
ALDP	652-667	FALLIIRPPLSLWYH
HPMP70	618-633	ITLFIIRKSLWYH
RPMP70	618-633	ITLFIIRKSLWYH
MDR1a	580-595	RTTVIARRLSTVRAH
MDR1b	1225-1240	RTCVIARRLSTVRAH
CFTRa	598-613	KTRRLCLCEHMEHLKA
CFTRb	1395-1410	CTVLCCEHMEHLKA
PSF1	754-769	RSLLIQLQLSLVQA
PSF2	671-686	RTLLIQLQLSLVQA
MALK	185-200	RTMYRQDQVEAMTL

The hydrophilic carboxy-terminal region of the ALD protein shows 56% identity over 210 amino acids to the corresponding region of 70 K PMP. The two characteristic nucleotide-binding consensus sequences²² are almost identical between the two proteins (underlined in Fig. 4b).

When northern blots of poly(A)⁺ RNAs from human tissues were hybridized to probes Ex13 or Ex3, a transcript of 4.2 kb was detected in heart, placenta, lung, liver, skeletal muscle, testis (not shown), pancreas and, to a lesser extent, in brain and kidney (Fig. 3b). The expression of the 4.2-kb transcript was very low in adult brain but more marked in 21-week fetal brain (not shown). A second transcript of 6.8 kb was detected in heart and skeletal muscle, whereas a third transcript of 2.75 kb was detected in muscle and liver. These additional transcripts could arise from alternative processing or the use of multiple polyadenylation sites. The sequence shown in Fig. 2 corresponds to 4.2-kb messenger RNA, this being the only species detected in HeLa cells.

We have identified the putative ALD gene in the distal part of Xq28 which has deletions in one or several exons in 6 of 85 independent ALD patients. Some of these deletions are small and non-overlapping, thus strengthening the conclusion that this gene is indeed involved in ALD. Although the gene coding for VLCFA-CoA synthetase was considered as a candidate gene for ALD, a recently cloned rat gene for long-chain (C₁₂-C₁₆) acyl-CoA synthetase²³ failed to detect homologous sequences on the X chromosome¹⁵. The putative ALD gene which we have now characterized shows no homology to this latter sequence, or to the three other enzymes involved in peroxisomal β -oxidation. Surprisingly, a very significant sequence identity was found

FIG. 4 a, Sequence alignment of ALD protein and human 70K PMP. Amino-acid identities are indicated by two dots and conservative changes by one. Sequence similarities were established with the FASTP program³⁶. b, Alignment of the C-terminal amino-acid sequence of ALD protein with sequences of human and rat 70K PMP and other members of the ATP-binding protein superfamily. Identity of amino acids between ALD protein and the other proteins are boxed in black and conservative amino-acid changes are boxed in grey. MDR1 is the human multidrug-resistant gene product²⁴; CFTR is the cystic fibrosis transmembrane conductance regulator²⁵; PSF1 and PSF2 are two peptide transporters^{26,27}; MALK is a bacterial protein necessary for maltose transport²⁷. Because of internal duplication in MDR1 and CFTR, the two ATP-binding regions (designated a and b) are represented. The two ATP-binding site motifs (NBF)²² are underlined and the 12-amino-acid segment that is highly conserved in other ABC proteins is indicated by asterisks.

with human and rat 70K PMP. Two putative domains could be identified by hydropathy analysis (not shown): an amino-terminal hydrophobic region, which presumably contains six transmembrane segments, and a hydrophilic region containing ATP-binding motifs with striking identity to the ATP-binding region of the human 70K PMP. This sequence is well conserved in the 'ATP-binding cassette' (ABC) family of transporters²⁹, which includes the multidrug-resistant gene product²⁴, the cystic fibrosis transmembrane conductance regulator²⁵, and the products of the *PSF1* and *PSF2* genes, which encode peptide transporters and map in the class II region of the human MHC complex^{26,27} (Fig. 4b). The conserved region includes the A and B consensus sequences (A: G-X4-G-K-T-X6-I/V; B: R/K-X3-G-X3-L(hydrophobic)-4-D) of a nucleotide-binding fold, plus a 12-amino-acid motif highly conserved in ABC proteins (Fig. 4b)²⁸.

ALD protein must therefore be a member of this superfamily of ABC transporters, which are also involved in transport of proteins, amino acids, inorganic ions and peptides in prokaryotes and eukaryotes²⁹. Although the predicted sequence of ALD protein shows significant identity to 70K PMP, no homology was found to 35K PMP or to other PMPs required for peroxisome biogenesis in yeast³⁰⁻³². Although ALD was initially thought to involve a deficiency in peroxisomal VLCFA-CoA synthetase, the predicted sequence of the putative ALD protein rather suggests a protein involved in transport of VLCFA-CoA synthetase into the peroxisomal membrane or a protein that is functionally associated with the VLCFA-CoA synthetase in the peroxisomal membrane. The translocation of acyl-CoA oxidase, the next enzyme or the peroxisomal β -oxidation pathway, requires ATP hydrolysis³³, whereas the transport of VLCFA across the peroxisomal membrane does not, and neither is it impaired in peroxisomes from ALD patients³⁴. Further studies are required to establish a peroxisomal localization for the ALD protein and to determine its role.

Expression of ALD protein was observed in every tissue tested, but the relationship between ALD protein expression and the abundance of peroxisomes in tissues may not be straightforward. Peroxisomes are particularly abundant in liver and kidney, having an average diameter of 0.2–1 μ M. In other tissues, including the brain and fibroblasts, they are less abundant and smaller (0.05–0.2 μ M)³⁹. This abundance and size difference may reflect a distinct membrane and matrix protein compositions of peroxisomes in different tissues. Although ALD is associated with a defective oxidation of VLCFA, this metabolic defect is mainly expressed in brain and adrenal tissues¹. At present it remains unresolved why brain and adrenal glands should be particularly sensitive to the accumulation of VLCFA.

ALD is characterized by a striking variation in clinical phenotype¹. In family Ma, an identical deletion was found in two sibs, a boy who developed cerebral ALD at 8 years, and his brother who developed only very mild adrenal insufficiency at 13 years. Furthermore, deletions were associated with the adult form (patients O and R) as well as with the severe childhood form (patients B and L). These differences suggest that the phenotypic variability of ALD is probably due to secondary factors (possibly immunological¹) or to the influence of still unidentified modifier genes³⁸. □

11. Nathans, J., Thomas, D. & Hogness, D. *Science* **232**, 193–202 (1986).
12. Sack, G. H., Raven, B. & Moser, H. W. *Am. J. hum. Genet.* **44**, 794–798 (1989).
13. Aubourg, P., Sack, G. H. & Moser, H. W. *Am. J. hum. Genet.* **42**, 408–413 (1988).
14. Aubourg, P. et al. *Am. J. hum. Genet.* **46**, 459–469 (1990).
15. Feil, R. et al. *Am. J. hum. Genet.* **49**, 1361–1371 (1991).
16. Kioschis, P. et al. *Cytogenet. Cell Genet.* **58**, 2070 (1991).
17. Kamijo, K., Taketani, S., Yokota, S., Osumi, T. & Hashimoto, T. *J. biol. Chem.* **265**, 4534–4540 (1990).
18. Kamijo, K., Kamijo, T., Ueno, I., Osumi, T. & Hashimoto, T. *Biochim. biophys. Acta* **1129**, 323–327 (1992).
19. Xiao, J. H., Davidson, I., Matthes, H., Garnier, J. M. & Chambon, P. *Cell* **65**, 551–568 (1991).
20. Kozak, M. *Nucleic Acids Res.* **15**, 8125–8148 (1987).
21. Gärtner, J., Moser, H. & Valle, D. *Nature Genet.* **1**, 16–23 (1992).
22. Walker, J. E., Saraste, M., Runswick, M. J. & Gay, N. J. *EMBO J.* **8**, 945–951 (1989).
23. Suzuki, H. et al. *J. biol. Chem.* **265**, 8681–8685 (1990).
24. Roninson, I. B. *Molecular and Cellular Biology of Multidrug Resistance in Tumor cells* (Plenum, New York, 1991).
25. Riordan, J. R. et al. *Science* **245**, 1066–1080 (1989).
26. Trowsdale, J. et al. *Nature* **348**, 741–744 (1990).
27. Bahram, S., Arnold, D., Bresnahan, M., Strominger, J. L. & Spies, T. *Proc. natn. Acad. Sci. U.S.A.* **88**, 10094–10098 (1991).
28. Cutting, G. R. et al. *Nature* **346**, 366–369 (1990).
29. Higgins, C. F. et al. *J. bioenerg. Biomembr.* **22**, 571–593 (1990).
30. Shimozawa, N. et al. *Science* **255**, 1132–1134 (1992).
31. Erdmann, R. et al. *Cell* **64**, 499–510 (1991).
32. Höfelf, J., Veenhuis, M. & Kunau, W. H. *J. cell Biol.* **114**, 1167–1178 (1992).
33. Imanaka, T., Small, G. M. & Lazarow, P. B. *J. cell Biol.* **105**, 2915–2922 (1987).
34. Singh, I., Lazo, O., Dhaunsi, G. S. & Contreras, M. J. *J. biol. Chem.* **267**, 13306–13313 (1992).
35. Sambrook, J., Fritsch, E. F. & Maniatis, T. *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Press, New York, 1989).
36. Lipman, D. J. & Pearson, W. R. *Science* **227**, 1435–1441 (1985).
37. Gilson, E., Nikaido, H. & Hofnung, M. *Nucleic Acids Res.* **10**, 7449–7458 (1982).
38. Moser, H. W. et al. *J. inher. metab. Dis.* **15**, 645–664 (1992).
39. Lazarow, P. B. & Moser, H. W. In *The Metabolic Basis of Inherited Diseases* (eds Scriver, C. R., Beaudet, A. L., Sly, W. S. & Valle, D.) 1479–1509 (McGraw-Hill, New York, 1989).

ACKNOWLEDGEMENTS. We thank S. Guidoux, C. Kretz and J. Lopez for technical assistance; J. M. Garnier for the HeLa cDNA library; S. Vicaire for DNA sequencing; I. Colas, P. Hamman and F. Ruffenach for oligonucleotide synthesis; F. Rocchiccioli for biochemical diagnosis; and P. F. Boughnères for encouragement. This work was supported by the Association Française contre les Myopathies, the Association Française contre l'Adréno-leucodystrophie, the Ministère de la Recherche et de la Technologie, the Deutsche Forschungsgemeinschaft and the European Communities, La Fondation Electricité de France, the United Leukodystrophy Foundation, the Centre Hospitalier Universitaire de Strasbourg, and Faculté Cochin.

A secreted protein kinase of *Yersinia pseudotuberculosis* is an indispensable virulence determinant

Edouard E. Galyov*, Sebastian Håkansson*, Åke Forsberg† & Hans Wolf-Watz*‡

* Department of Cell and Molecular Biology, University of Umeå, S-90187 Umeå, Sweden

† Department of Microbiology, National Defence Research Establishment, S-90182 Umeå, Sweden

PHOSPHORYLATION of proteins catalysed by protein kinases is associated with central functions in growth and proliferation of the eukaryotic cell, and kinases are particularly important in the signal transduction pathways^{1,2}. Enterobacterial protein kinases are structurally and functionally different from eukaryotic protein kinases^{3,4}, and no prokaryotic kinase has so far been described implicating a direct role for this activity in virulence. Virulent *Yersinia* possess a common virulence plasmid that encodes a number of secreted proteins (Yops)^{5–7}, of which YopH has protein-tyrosine phosphatase activity with a key function in the block of phagocytosis by the pathogen^{8–10}. Here we report that the virulence plasmid of *Yersinia pseudotuberculosis* encodes a secreted protein kinase (YpkA) with extensive homology to eukaryotic Ser/Thr protein kinases^{3,11}. Specific mutants of *ypkA* resulted in avirulent strains. Thus, YpkA is, to our knowledge, the first reported prokaryotic secreted protein kinase involved in pathogenicity, presumably by interfering with the signal transduction pathways of the target cell.

‡ To whom correspondence should be addressed.

Received 1 September 1992; accepted 11 January 1993.

1. Moser, H. W. & Moser, A. In *The Metabolic Basis of Inherited Disease* (eds Scriver, C. R., Beaudet, A. L., Sly, W. S. & Valle, D.) 1511–1532 (McGraw-Hill, New York, 1989).
2. Singh, I., Moser, A. E., Goldfischer, S. & Moser, H. W. *Proc. natn. Acad. Sci. U.S.A.* **81**, 4203–4207 (1984).
3. Hashmi, M., Stanley, W. & Singh, I. *FEBS Lett.* **196**, 247–250 (1986).
4. Lazo, O. et al. *Proc. natn. Acad. Sci. U.S.A.* **85**, 7647–7651 (1988).
5. Wanders, R. J. A. et al. *J. inher. metab. Dis.* **11**, 173–177 (1988).
6. Überbacher, E. C. & Mural, R. *Proc. natn. Acad. Sci. U.S.A.* **88**, 11261–11265 (1992).
7. Fearon, E. R. et al. *Science* **247**, 49–56 (1990).
8. Migeon, B. R. et al. *Proc. natn. Acad. Sci. U.S.A.* **78**, 5066–5070 (1981).
9. Boué, J. et al. *Hum. Genet.* **69**, 272–274 (1985).
10. Aubourg, P., Sack, G. H., Meyers, D. A., Lease, J. J. & Moser, H. W. *Ann. Neurol.* **21**, 349–352 (1987).

The genus *Yersinia* includes three pathogenic species, *Y. pestis*, *Y. pseudotuberculosis* and *Y. enterocolitica*, which all harbour a common virulence plasmid of about 70 kilobases (kb) (Fig. 1a). In response to the extracellular stimuli, calcium and temperature, a number of plasmid-encoded proteins (Yops) are secreted to the culture supernatant (Fig. 1b)⁵⁻⁷. Three of these, YopE, YopH and YopM, are essential virulence determinants¹²⁻¹⁴. YopH, having structural similarities to eukaryotic protein tyrosine phosphatases, was recently found to have a protein-tyrosine phosphatase activity and thus may cause dephosphorylation of host proteins^{9,10}.

That *Yersinia* secrete a protein-tyrosine phosphatase (YopH) suggested that the pathogens, in addition, express a secreted protein kinase. Therefore, proteins of *Y. pseudotuberculosis* secreted to the culture supernatant were concentrated by ultrafiltration and were used in a protein kinase assay with [γ -³²P]ATP as the phosphate donor. In parallel experiments, [α -³²P]ATP was used to discriminate between phosphotransferase activity and stable association of ATP with the proteins. Labelled proteins in the size range of 27 to 49K, including a predominantly labelled protein of 27K, were detected with equal intensity using either [γ -³²P]ATP or [α -³²P]ATP (data not shown). But, two proteins of about 80K and 49K were specifically labelled when using [γ -³²P]ATP, showing that there was a protein kinase activity in this preparation (Fig. 1c). To identify the gene responsible for this possible kinase activity, one of the regions of the virulence plasmid pIB1 of *Y. pseudotuberculosis*, which had not been explored earlier, was cloned and sequenced (Fig. 1a). One open reading frame was found that encoded a protein of 81.7K (Figs 1a and 2), which

showed noticeable homology to Ser/Thr protein kinases of eukaryotic origin (the PSK family)^{3,11} (Fig. 3). The protein was therefore denoted YpkA (*Yersinia protein kinase*). YpkA contains a putative catalytic domain, located at the amino terminus of the protein which contains structural motifs, subdomains I to XI, common to all eukaryotic protein kinases¹¹ (Fig. 2). The consensus sequence Asp-Ile-Lys-Pro-Gly-Asn of subdomain VI indicates that YpkA has a specificity for serine and/or threonine¹¹. But the presence of a glycine residue in this sequence may also indicate an additional tyrosine kinase specificity for YpkA¹⁵ (Fig. 3).

To study expression, activity and possible involvement in virulence of YpkA, two *ypkA* mutants were constructed. YPIII/pIB43 is an insertion mutant in which the YpkA protein has been disrupted downstream of the catalytic domain at amino-acid residue 548 (Fig. 2). The strain YPIII/pIB44 is an in frame deletion mutant of *ypkA* in which 546 bp, corresponding to amino acids 207 to 388 of YpkA, were deleted. This deletion removes a major portion of the catalytic domain (Fig. 2). When secreted proteins were analysed, the wild-type strain expressed a protein with an apparent M_r of 80K, whereas the two mutants expressed the expected truncated forms of the mutated *ypkA* (Fig. 1b).

As expected, the 80K and 49K proteins were not detected when the *ypkA* mutant YPIII/pIB44 was assayed for protein kinase activity (Fig. 1c). When the YPIII/pIB43 mutant was analysed, a few poorly labelled proteins could be detected which may be explained by a residual activity of the truncated YpkA protein (Fig. 1c). These results strongly suggest that the kinase activity is associated with the YpkA protein and that YpkA can

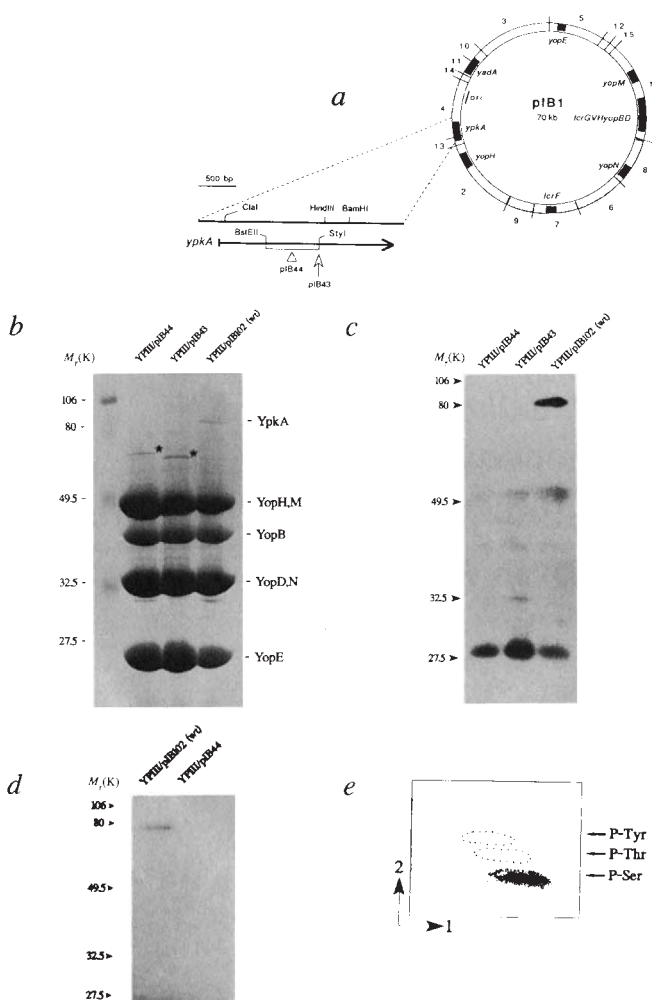


FIG. 1 Analysis of the *ypkA* gene by restriction mapping, SDS-PAGE, kinase assay and phosphoamino acid analysis. *a*, Restriction map of pIB1 of *Y. pseudotuberculosis*. The numbers refer to the commonly used *Bam*HI map of pIB1¹⁷. The enlarged region shows the position and partial restriction map of the *ypkA* gene. The positions of the constructed *ypkA* mutations are indicated and the names of the plasmids carrying these mutations are shown below. *b*, SDS-PAGE protein analysis shows the secreted proteins of the wild-type YPIII/pIB102 and the two *ypkA* mutants, YPIII/pIB43 and YPIII/pIB44. The position of YpkA in the wild-type strain is shown and the stars indicate the positions of the YpkA truncated forms from the two mutants. *c*, The protein phosphorylation of the secreted proteins from the wild-type and the two mutants of *ypkA* is shown. *d*, Autophosphorylation of immunoprecipitated YpkA. *e*, Phosphoamino acid analysis of the autophosphorylated YpkA protein.

METHODS. The 2.5-kb DNA fragment indicated was sequenced using the method of Sanger. A 1.5-kb *Bam*HI-*Cla*I fragment included in *Bam*HI fragment 4 was cloned into the pBluescript vector. The hybrid plasmid was cut with *Bst*EII and *Sty*I, the ends were filled in by the use of the Klenow enzyme and the plasmid was religated. This resulted in a 546-bp (amino-acid positions 207 to 388) in-frame deletion within the cloned portion of *ypkA*. The deletion was confirmed by DNA sequence analysis. The fragment carrying the deletion was first integrated into the pIB102 virulence plasmid by a single recombination step using the suicide vector pNQ705¹⁸ as vehicle. This integration resulted in the mutant strain YPIII/pIB43. The in-frame deletion of the *ypkA* gene on the virulence plasmid was obtained by excision of the suicide vector. Protein kinase assay: bacteria were grown in a defined liquid TMH medium with conditions allowing maximal expression of plasmid-encoded proteins¹⁴. Proteins from the culture supernatant were concentrated 25-fold by ultrafiltration¹⁴. During ultrafiltration, the medium was changed to a buffer containing: 10 mM Tris-HCl pH 7.5, 50 mM KCl, 10 mM MgCl₂. To 200 μ l of the sample, 20 μ l of [γ -³²P]ATP (or [α -³²P]ATP) was added and the sample was incubated for 30 min at 37 °C. The proteins were precipitated by 10% trichloroacetic acid and washed with acetone. The samples were solubilized in SDS-sample buffer containing 5 mM ATP and subjected to SDS-PAGE followed by autoradiography. Immunoprecipitation was done as described¹⁹ using a monospecific rabbit anti-YpkA antiserum. Immunization and purification of the monospecific YpkA antiserum was performed as described for the YopN protein²⁰. The 80K phosphorylated protein was excised and analysed for phosphoamino acids¹⁶.

autophosphorylate. To investigate this further, immunoprecipitates of the wild-type and the mutant YpkA proteins, obtained by using monospecific anti-YpkA antibodies, were subjected to the kinase assay. One labelled 80K protein was detected in the wild-type, whereas no proteins were labelled when the mutant YPIII/pIB44 was assayed (Fig. 1d). Thus, YpkA possesses autophosphorylating activity. Phosphoamino acid analysis¹⁶ of the labelled 80K protein revealed that this protein contained phosphorylated serine residues (Fig. 1e).

To test the effect of the mutations on virulence, BALB/c mice were challenged orally with the wild-type and the mutant strains. No lethal infection was observed with the mutant strains, whereas all mice died when challenged with the wild-type strain (Table 1). Thus, YpkA is an indispensable virulence determinant of *Y. pseudotuberculosis*.

```

1 - M K S Y K I M G T M P P S I S L A K A H E R I S Q H W Q N P Y G E L N I G O K R Y R I D N Q V L R L N P H S G F S L F
61 - R E G Y G K I P S Q K M F N P S I A R N L T D T L H A A Q K T T S Q E L R S D I P N A L S N L F G A K P Q T E L P L G W
121 - K G E P L S G A P D L R G M R Y A E T D K F A R Q E S H I S I E T K D K Q L V A K I E R S I A R G H L F A E L E A Y
181 - K H I Y K T A G E N P K L A N V H M A V Y P G N K R E S A L M D Y D W K C S D T L R T L A D S W K Q K I N S
241 - E A Y W O T I K F I A R L L D Y T N H L A K A Q V Y H N D I K P O N Y V P D R A S O E P Y V I D L G L H S R S O E Q P
301 - K G F T E S F A P E L G V N L G A S E K S D V F L V Y S T L L N C I E G F E K N P E I K P N Q L R F I T S E P A H
361 - Y M D E N G Y P I R R P G I A G V E T A Y T R P I T D I L G Y S A D S R P S N E A R L H E F L S D G T I D E E S A K Q
421 - I L K D T L T G E M S P L S T D V R R I T F K K L E L S D L L T H S S A A T Q L D M G G V I S D I D T M L V A L
481 - D K A E R E G O V D K D Q L K S P N S I L K T Y R V I E D Y Y K G R E O D T K N S S T E V S P Y H R S N P M L S I V E
541 - P S L Q R I Q K L D T O S P S D I G S L V R A N K H L E T L L E V L T L S Q Q Q P V S S E T Y P L N L A E A
601 - K I T L S Q L L N T Q Q Q E S A K A Q L S I L I N S O S W A D V A R Q S L Q R F D S T R P Y K F G T E Q Y T A I
661 - H R Q M A A H A A I T L Q V S E P T D M N E N P T V D S I P L I Q L O R S S L M D H L V E Q R E K L R E L T T I
721 - A R K L N R L E R E W M

```

FIG. 3 Multiple alignment of subdomains six to eight of the catalytic domains of YpkA and some representative eukaryotic protein kinases³. The subdomains are indicated with Roman numerals according to Hanks *et al.*¹¹. Conserved amino-acid residues of YpkA and the different eukaryotic protein kinases are boxed. The eukaryotic protein kinases are divided into two groups on the basis of their specificity. PSKs indicate serine/threonine specificity and PTKs indicate tyrosine specificity^{3,11}.

Plasmid-encoded proteins having M_r s similar to YpkA were also observed when strains of *Y. pestis* and *Y. enterocolitica* were analysed^{6,7}. Owing to the high homology between the virulence plasmid of all pathogenic *Yersinia*, it should come as no surprise if all strains of pathogenic *Yersinia* possess the protein kinase essential for virulence. The YpkA protein is the third virulence plasmid-encoded protein of *Yersinia* that exhibits high homology with eukaryotic counterparts. This favours a hypothesis that the three *Yersinia* genes were recruited from a eukaryotic source through horizontal gene transfer.

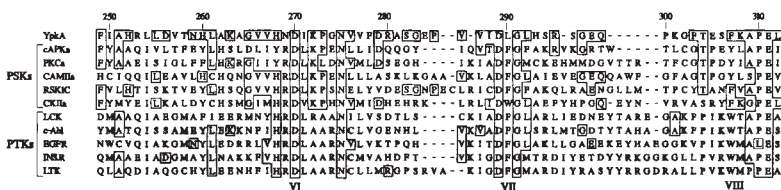
Although its cellular function and molecular target are unknown, YpkA probably has a pronounced effect on the level of serine and/or threonine phosphate in the target cells. Increased understanding of the mechanisms used by the pathogens should improve our knowledge of both microbial pathogenesis as well as the mechanisms of signal transduction of eukaryotic cells. □

TABLE 1 Virulence test of the wild-type and ypkA mutant strains

Strains	Number of mice surviving oral infection	
	Experiment 1	Experiment 2
YPIII/pIB102	0/3	0/3
YPIII/pIB43	3/3	3/3
YPIII/pIB44		3/3

BALB/c mice were infected orally as described earlier¹⁷. The mice were challenged with the same dose of bacteria (5×10^8 bacteria per ml of drinking water). The progress of each infection was monitored for 14 days. All mice were dead five days after challenge with the wild-type strain, YPIII/pIB102.

FIG. 2 The deduced amino-acid sequence (single-letter code) of the YpkA protein kinase translated from the corresponding DNA sequence. The conserved subdomains of YpkA corresponding to the characterised subdomains of eukaryotic protein kinases are numbered with Roman numerals (I–XI)¹¹. The amino acids deleted in mutant YPIII/pIB44 are marked by arrows marked with triangles. The point of interruption of the YpkA reading frame by the insertion in mutant YPIII/pIB43 is indicated by a vertical arrow. The borders of each subdomain are not definite because of the distant relationship of YpkA to other kinases. The DNA sequence was submitted to the EMBL Data library with the accession number X69439.



- Bölin, I. & Wolf-Watz, H. *Molec. Microbiol.* **2**, 237–245 (1988).
- Rimpiläinen, M., Forsberg, Å. & Wolf-Watz, H. *J. Bact.* **174**, 3355–3363 (1992).
- Forsberg, Å., Bölin, I., Norlander, L. & Wolf-Watz, H. *Microbiol. Pathogen.* **2**, 123–137 (1987).
- Forsberg, Å., Viitanen, A.-M., Skurnik, M. & Wolf-Watz, H. *Molec. Microbiol.* **5**, 977–986 (1991).

ACKNOWLEDGEMENTS. This work was supported by grants from the Swedish Medical Research Council, the Swedish Natural Science Research Council and the Kempe foundation. E.E.G. is a postdoctoral fellow supported by the Royal Swedish Academy of Science and the Institute of Immunology, Moscow region 142380, Russia.

Extracellular domain of the boss transmembrane ligand acts as an antagonist of the sev receptor

A. C. Hart*, H. Krämer* & S. L. Zipursky†

Howard Hughes Medical Institute, Department of Biological Chemistry, The Molecular Biology Institute, University of California, Los Angeles, Los Angeles, California 90024, USA

THE fate of the R7 photoreceptor cell in the *Drosophila* compound eye is established by a specific inductive interaction between the R8 photoreceptor neuron and the R7 precursor cell¹. This induction is mediated by two cell-surface proteins: the ligand, bride of sevenless² (boss), and sevenless (sev), a tyrosine-kinase receptor^{3–5}. The structure of boss is unique for a ligand of a tyrosine-kinase receptor. It contains a large extracellular domain, seven transmem-

Received 8 October; accepted 18 December 1992.

- Nurse, P. *Nature* **344**, 503–508 (1990).
- Pelech, S. L. & Sanghera, J. S. *Trends biochem. Sci.* **17**, 233–238 (1992).
- Hanks, S. K. *Curr. Opin. Struct. Biol.* **1**, 369–383 (1991).
- Cozzzone, A. J. *Rev. Microbiol.* **4**, 97–125 (1988).
- Bölin, I., Portnoy, D. A. & Wolf-Watz, H. *Infect. Immun.* **48**, 234–240 (1985).
- Cornelis, G., Vanooteghem, J.-C. & Sluiters, C. *Microbiol. Pathogen.* **2**, 367–379 (1987).
- Straley, S. C. & Bowmer, W. S. *Infect. Immun.* **51**, 445–454 (1986).
- Rosqvist, R., Bölin, I. & Wolf-Watz, H. *Infect. Immun.* **56**, 2139–2143 (1988).
- Guan, K. L. & Dixon, J. E. *Science* **249**, 553–556 (1990).
- Clemens, J. C., Guan, K., Bliska, J. B., Falkow, S. & Dixon, J. E. *Molec. Microbiol.* **5**, 2617–2620 (1991).
- Hanks, S. K., Quinn, A. M. & Hunter, T. *Science* **241**, 42–52 (1988).
- Rosqvist, R., Forsberg, Å., Rimpiläinen, M., Bergman, T. & Wolf-Watz, H. *Molec. Microbiol.* **4**, 657–667 (1990).
- Leung, K. Y., Reisner, B. S. & Straley, S. C. *Infect. Immun.* **58**, 3262–3271 (1991).
- Rosqvist, R., Forsberg, Å. & Wolf-Watz, H. *Infect. Immun.* **59**, 4562–4569 (1991).
- Lindberg, R. A., Quinn, A. M. & Hunter, T. *Trends biochem. Sci.* **17**, 114–119 (1992).
- Cooper, J. A., Sefton, B. M. & Hunter, T. *Meth. Enzym.* **99**, 387–402 (1983).

* Present addresses: Department of Molecular Biology, Wellman 8, Massachusetts General Hospital, Boston, Massachusetts 02114, USA (A.C.H.); Department of Cell Biology and Neuroscience, University of Texas, Southwestern Medical Center, Dallas, Texas 75235, USA (H.K.).

† To whom correspondence should be addressed.

brane segments, and a carboxy-terminal cytoplasmic tail^{6,7}. Here we report that: (1) boss activates tyrosine phosphorylation of the sev receptor; (2) the seven transmembrane domain of boss is necessary for its function; and (3) a soluble form of boss acts as an antagonist of the sev receptor both *in vivo* and *in vitro*.

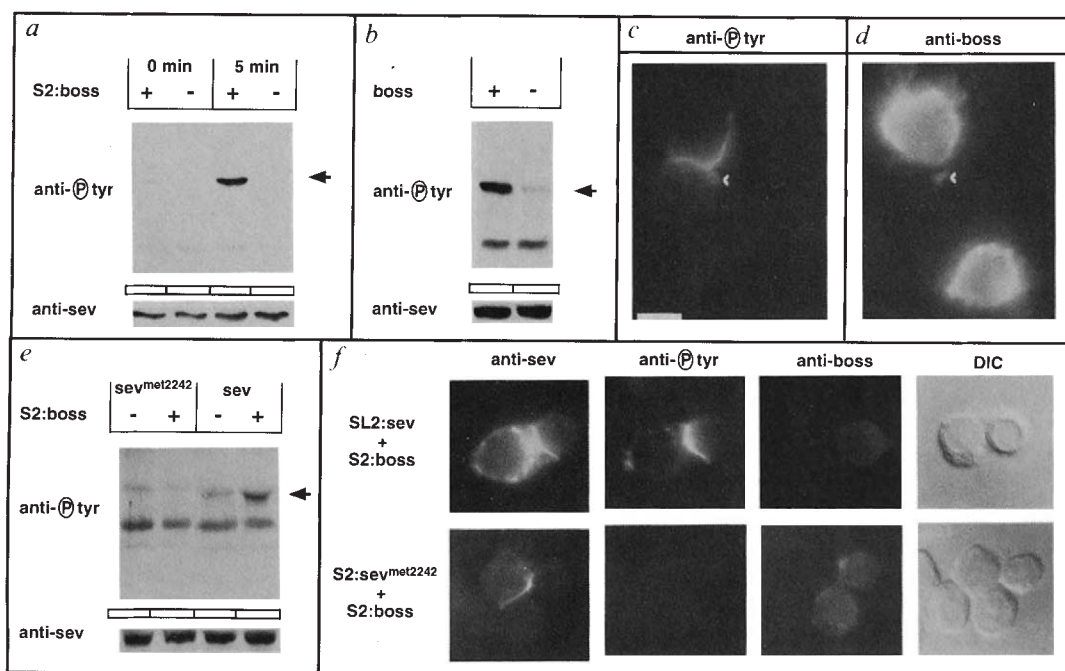
Binding of boss-expressing cells (S2:boss) to sev-expressing cells (SL2-sev) resulted in a rapid increase in tyrosine phosphorylation of the sev receptor (Fig. 1a). Membrane preparations of S2:boss cells also induced sev tyrosine phosphorylation (Fig. 1b). Neither S2 cells nor S2 membrane preparations altered the level of phosphotyrosine on sev, providing evidence for specificity. In addition, S2:boss cells did not stimulate tyrosine phosphorylation in cells expressing the sev^{met2242} protein (Fig. 1e). This mutant protein binds to boss², but is inactive as a tyrosine kinase owing to a substitution of methionine for lysine 2,242 (ref. 5) thought to be critical for the phosphotransfer reaction. These data are consistent with activation of the sev tyrosine-kinase receptor by binding to boss.

The cellular distribution of anti-phosphotyrosine immunoreactivity in SL2-sev cells in contact with S2:boss cells was also assessed. Co-aggregates of S2:boss and SL2-sev cells were triple-stained with antibodies to boss, sev and phosphotyrosine.

Intense phosphotyrosine staining was observed at regions of contact between the boss and sev-expressing cells (Fig. 1c,f). Boss-containing vesicular structures within the SL2-sev cells often colocalized with intense phosphotyrosine immunoreactivity (Fig. 1c,d). But vesicular anti-phosphotyrosine staining was also seen in SL2-sev cells incubated alone (data not shown). Cells expressing the sev^{met2242} protein did not show an increase in phosphotyrosine at regions of contact with boss-expressing cells (Fig. 1f). In both SL2-sev cells and S2 cells expressing sev^{met2242}, the sev protein is clustered at regions of contact with S2:boss cells. We conclude that the increase in tyrosine phosphorylation in the region of contact between the S2:boss and SL2-sev cells is due to local activation of the sev tyrosine-kinase receptor by boss.

Soluble forms of other transmembrane ligands (such as TGF- α (refs 8, 9) and *Steel* (refs 10–14) bind with high affinity to receptor tyrosine kinases and stimulate tyrosine-kinase activity. To test whether the boss extracellular domain is sufficient to interact with the sev receptor, we examined the activity of a secreted form of boss, EXboss, which contains only the large extracellular amino-terminal domain of boss (Fig. 2a). The EXboss protein was produced in S2 cells under the control of

FIG. 1 Tyrosine phosphorylation of sev is stimulated by boss. **a**, **b**, **e** Tyrosine phosphorylation of the immunoprecipitated sev receptor was assessed by western blots with an anti-phosphotyrosine antibody, mAb4G10. Arrows indicate the C-terminal fragment of the sev receptor containing the tyrosine kinase domain. Parallel blots were probed with an anti-sev antibody, mAb150C3 (ref. 15), to assess the amount of immunoprecipitated sev protein. The identities of the other bands visible in the westerns are not known. **a**, Binding of S2:boss cells (+), but not S2 cells (–), to SL2-sev cells stimulates tyrosine phosphorylation of sev. **b**, Plasma membranes (15 μ g) purified from S2:boss cells (+) but not S2 cells (–), induced sev tyrosine phosphorylation. **c**,



d, Immunofluorescence staining of bound sev- and boss-expressing cells. S2:boss cells bound to SL2-sev cells were stained with antibodies to phosphotyrosine (anti-P-tyr) and boss (anti-boss). Phosphotyrosine immunoreactivity was concentrated at the interface between S2:boss and SL2-sev cells. Vesicles within the SL2-sev cell were detected that contained both phosphotyrosine (**c**) and boss immunoreactivity (**d**; arrowheads). **e**, Tyrosine phosphorylation requires sev kinase activity. Binding of S2:boss cells to S2 cells expressing the sev^{met2242} protein (S2:sev^{met2242} cells) does not increase sev tyrosine phosphorylation. The sev^{met2242} protein⁵ contains a methionine in place of a lysine which is required for the phosphotransfer reaction. **f**, Concentration of phosphotyrosine at the interface between sev- and boss-expressing cells requires the catalytic activity of sev. Localization of phosphotyrosine in SL2-sev (top) or S2:sev^{met2242} (bottom) cells in contact with S2:boss cells was compared. Cells were triple stained for phosphotyrosine, boss and sev as described below. Capping of the sev^{met2242} receptor in the membrane apposing the boss-expressing cells indicates that the mutant sev receptor binds to the boss ligand. Differential interference contrast (DIC) was used in the photos indicated. Scale bar represents 6 μ m in **c** and **d**, 8 μ m in **f**.

METHODS. **a**, Cell lines: S2:boss (ref. 2) and SL2-sev (ref. 4) were previously described. The S2:sev^{met2242} cells were generated by the co-transfection² of pPC4 (conferring α -amanitin resistance) and pKB174 (a gift of E. Hafen) which contains the sev^{met2242} gene under the control of the *actin5C* pro-

motor. The S2:sev^{met2242} cells were not cloned, rather a pool of transfected cells was used after 3 weeks of selection with α -amanitin. **b**, Phosphotyrosine assays. To measure tyrosine phosphorylation of sev, 2×10^6 SL2-sev cells were mixed with 4×10^6 S2:boss cells and forced into contact by pelleting for 1 min at 300g. After 5 min at 25 °C, cell pellets were solubilized in 1 ml 1% Triton X-100, 150 mM NaCl, 100 mM Tris-HCl, pH8, 50 μ g ml⁻¹ leupeptide, 50 μ g ml⁻¹ palpain, 1 mM PMSF and 200 μ M Na₃VO₄. Sev protein was immunoprecipitated using 1 μ l anti-sev antiserum G15 (ref. 20) and Protein G-Sepharose. Western blots were done as described². The size of the C-terminal sev fragment was estimated to be 60K using prestained molecular mass markers (BRL) as previously described². mAb4G10 (UBI) and mAb150C3 were used at a 1:2,000 dilution. To compensate for the higher level of sev expression in the SL2-sev line versus the S2:sev^{met2242} cells, 3×10^6 S2:sev^{met2242} cells were compared to 2×10^5 SL2-sev cells diluted 15 times with S2 cells. The sev cells were then mixed with 6×10^6 S2 or S2:boss cells and processed as above. **c**, Immunofluorescence microscopy. Immunohistochemical analysis of cells was done essentially as previously described² with the addition of 200 μ M Na₃VO₄ to all buffers. Primary antibodies were anti-boss CT2 (rabbit)⁷, anti-phosphotyrosine mAb4G10 (mouse, UBI) and anti-sev antiserum G14 (goat)²⁰. Donkey secondary antibodies (Jackson laboratories) were coupled to Texas Red (anti-rabbit), to FITC (anti-mouse) or to AMCA (anti-goat).

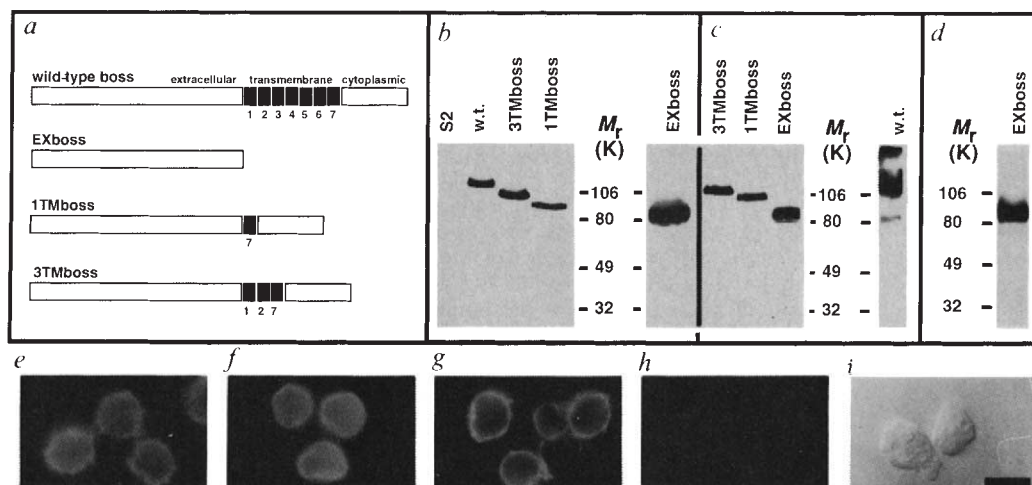
hsp70 promoter and was purified by lectin affinity chromatography (Fig. 2d). Radioiodinated preparations of EXboss failed to reveal high-affinity binding sites on SL2-sev cells. EXboss could not stimulate *sev* tyrosine kinase activity (Fig. 3b).

To assess whether the inability of EXboss to activate the *sev* receptor reflected a requirement for boss to be tethered to the

cell membrane, the activities of two mutant membrane-bound forms of boss were examined. The 1TMboss protein lacks transmembrane domains 1–6, whereas 3TMboss is deleted for transmembrane domains 3–6 (Fig. 2a). These mutant forms of boss were expressed in S2 cells at levels similar to wild type as estimated from western blots (Fig. 2b). The level of cell-surface

FIG. 2 Expression of mutant boss proteins. **a**, Schematic diagram comparing the structure of wild-type and mutant boss proteins used in this analysis. The transmembrane segments are numbered below the schematics. **b**, These proteins, stably expressed in S2 cell lines, migrate in SDS-PAGE gels as expected from the predicted size. Wild-type boss (w.t.), 3TMboss and 1TMboss were detected with pAbCT2 (ref. 7). EXboss is secreted into the medium by transfected S2 cells and detected with pAbED1 (ref. 7). **c**, Mutant proteins from third instar larvae and wild-type boss

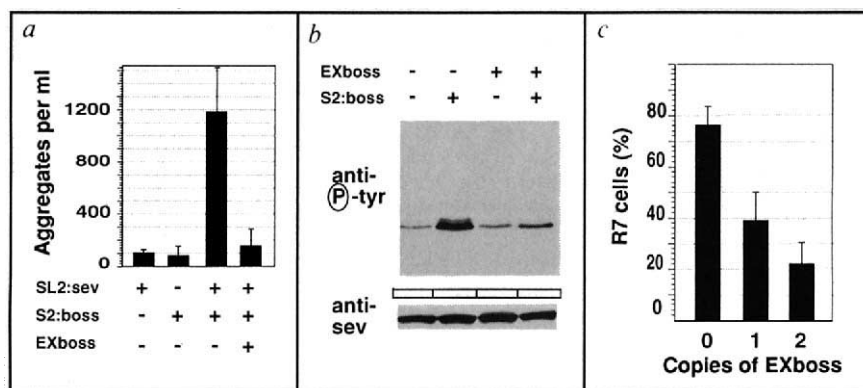
protein from adult head were expressed in transgenic animals using the *hsp70* promoter. Equal aliquots are loaded in each mutant protein lane and proteins were detected with pAbED1. **d**, Silver-stained gel containing 3 μ g affinity-purified EXboss protein used in *in vitro* assays. **e–i**, The orientation of 3TMboss expressed in S2 cells was assessed using mAbNT1 and pAbCT2, which recognize the N and C terminus of the protein, respectively⁷. Membrane permeabilization with 0.1% saponin reveals both N-terminal (**e**) and C-terminal epitopes (**f**). Without saponin, N-terminal (**g**) epitopes are accessible, but C-terminal (**h**) epitopes are not. **i**, Cells in **h** shown using differential interference contrast microscopy. The orientation of 1TMboss in S2 cells was confirmed using the same antibodies (data not shown). Scale bar, 8 μ m. **METHODS**. Mutant proteins were generated by internal deletions in the boss cDNA⁶. EXboss results from a ligation of the *EcoRV* site at amino acid 510, before transmembrane segment TM1, to the *SmaI* site of Bluescript changing amino acid 511 to Trp and adding Gly-Ile-His before translation termination. 1TMboss results from the ligation of the same *EcoRV* site to a *StuI* site in TM7, causing the deletion of amino acids 511–760 and the creation of a Ser at the junction. 3TMboss results from the ligation of the *AccI* site in TM3 with the *StuI* site in TM7, causing the deletion of amino acids 613–760 and the creation of a Ser. The deletion constructs were inserted into the



pHT4-*hsp70* expression vector for transformation as previously described¹⁷. Stable S2 cell lines and multiple fly lines were generated for all constructs. Homogenates from S2 cells were western blotted as previously described². The position of prestained molecular mass markers (BRL) is indicated. To prepare homogenates for western blotting, third instar larvae containing mutant protein constructs were incubated for 60 to 75 min at 37 °C, then for 60 min at 25 °C and processed as previously described for adult heads². The wild-type boss sample from adult head was processed as previously described². EXboss was purified by peanut agglutinin affinity chromatography. EXboss:S2 cells were heat-shocked for 30 min at 37 °C, then washed twice with serum-free Schneider's medium. EXboss was secreted for 3 h into serum-free medium. Filtered medium (100 ml) at a final concentration of 150 mM NaCl and 10 μ M MnCl₂ was incubated with 1 ml peanut agglutinin agarose (Vector Laboratories). The agarose was washed with 50 ml buffer containing 200 mM NaCl, 0.1 mM CaCl₂, 10 μ M MnCl₂, 100 mM Tris, pH7.3, and eluted with 3 ml of the same buffer including 0.2 M galactose. Using Centricon30 filters the protein was concentrated to 1 mg ml⁻¹ in PBS (10 mM sodium phosphate, pH7.2, 140 mM NaCl, 3 mM KCl). Staining of boss proteins expressed in S2 cells with mAbNT1 and pAbCT2 in the presence or absence of saponin was done as previously described⁷.

FIG. 3 EXboss acts as an antagonist of the *sev* receptor *in vitro* and *in vivo*. **a**, The aggregation of *sev*-expressing and boss-expressing S2 cells² is inhibited by the addition of 40 μ g ml⁻¹ soluble EXboss. **b**, The boss-dependent tyrosine phosphorylation of the *sev* receptor is inhibited by the addition of 100 μ g ml⁻¹ of soluble EXboss. Either S2:boss (+) or S2 cells (-) were mixed with SL2-sev cells and *sev* tyrosine phosphorylation was assessed as described in Fig. 1. **c**, In a genetic background where wild-type boss expression is limiting and the number of ommatidia containing R7 cells is ~80%, EXboss, present in one or two copies, antagonizes wild-type boss function. Seven to fourteen eyes were sectioned for each determination of mutant protein activity. An average of 81 ommatidia were counted per eye. Neither 1TMboss nor 3TMboss affect R7 development.

METHODS: Aggregation assays were done as previously described². To test inhibition of aggregation by EXboss, 40 μ g ml⁻¹ of purified EXboss was added to the SL2-sev cells immediately before the addition of boss-expressing cells. In *sev* tyrosine phosphorylation assays, 100 μ g ml⁻¹ EXboss and either 4 $\times$ 10⁶ S2:boss or S2 cells were added simultaneously to 2 $\times$ 10⁶ SL2-sev cells in 100 μ l. After 6 min, cells were lysed and processed as described in Fig. 1. The *D. melanogaster* boss rescue construct, G69, contains a 7.5 kb *XhoI*-*SalI* genomic fragment containing the entire



boss gene inserted into the *SalI* site of pDM23 (our unpublished observations). All boss⁺ third instar larvae containing one copy of the G69 construct were incubated from 60 to 75 min at 37 °C. Plastic sections (1.5 μ m) of adult eyes were scored for the presence of R7 cells as previously described¹⁷. The G69 boss rescue construct gives 5–15% rescue of R7 cells in heterozygous females when raised at 25 °C, but heat-shock treatment causes an increase to 80% rescue of R7 cells for more than 14 ommatidial rows. The cause of this increase is not known.

expression of wild type, 1TMboss and 3TMboss also appeared similar. Staining with a panel of antibodies revealed that, as for wild-type boss, the N- and C-terminal domains in these two mutant proteins were found on the extracellular and intracellular sides of the plasma membranes, respectively (Fig. 2*e-i*). Neither 1TMboss- nor 3TMboss-expressing cells aggregated with SL2-sev cells; forced contact by pelleting the mutant boss cells and SL2-sev cells together did not result in an increase in sev tyrosine phosphorylation (data not shown). These data suggest that the seven-transmembrane segment does more than simply anchor the N-terminal extracellular domain of boss to the plasma membrane.

The EXboss protein inhibited the interaction between boss and sev both *in vivo* and *in vitro*, indicating that this domain of boss interacts with sev. Purified EXboss at a concentration of $40 \mu\text{g ml}^{-1}$ decreased the aggregation of S2:boss and SL2-sev cells by more than 90% (Fig. 3*a*). In addition, EXboss inhibited the ability of boss-expressing cells to stimulate the sev tyrosine kinase (Fig. 3*b*). The ability of EXboss to inhibit the boss-sev

interaction *in vivo* was assessed in transgenic flies carrying one or two copies of EXboss expressed under the control of the *hsp70* promoter. Although EXboss did not interfere with the function of one copy of the *boss*⁺ gene on the third chromosome, EXboss inhibited R7 development under conditions in which boss is a limiting component in the inductive pathway. In a genetic background in which the boss phenotype is only partially rescued because of reduced expression of wild-type boss (that is, from a boss rescue line, G69; A.C.H. *et al.*, unpublished observations) expression of EXboss from one copy of the *hsp70* construct decreased the percentage of ommatidia containing R7 cells from $76 \pm 7\%$ to $39 \pm 10\%$. Two copies of the EXboss construct further reduced the rescue to $23 \pm 7\%$ (Fig. 3*c*). No other effects were observed in response to EXboss expression.

Immunohistochemical localization of EXboss in the developing eye disc indicates that inhibition of R7 development by EXboss was probably due to its direct binding of sev. Antibodies to sev reveal a dynamic pattern of expression^{15,16} in the developing eye disc. Sev is localized to the apical membranes of R1, R3, R4, R6 and R7, the so-called mystery cells and the cone cells. The same cellular pattern of apical localization was seen for the EXboss protein expressed under the control of the *hsp70* promoter (compare Fig. 4*a* with *b*). EXboss was also found in a perinuclear region of all cells in the developing disc in a pattern consistent with localization to the endoplasmic reticulum (data not shown). The apical localization of EXboss was not seen in a *sev*^{d2} background (Fig. 4*c*), whereas the basal staining was unchanged. Note that wild-type boss, expressed ectopically using the *hsp70* promoter, is apically localized independently of sev expression¹⁷. We conclude that EXboss is either secreted from cells in the developing disc and then bound by sev-expressing cells or EXboss associates intracellularly with sev, perhaps within the endoplasmic reticulum, followed by co-transport to the apical region of the cell.

Although 1TMboss and 3TMboss are targeted to the plasma membrane in transfected S2 cells (see Fig. 2), they accumulate in the perinuclear endoplasmic reticulum region of cells in the eye imaginal disc (data not shown). Thus, it is not possible rigorously to assess their ability to induce R7 development. But like EXboss, 1TMboss and 3TMboss are localized apically in sev-expressing cells and this apical localization is sev-dependent (data not shown). But in contrast to EXboss, mutant protein expression from one copy of 1TMboss and 3TMboss, or two copies of 1TMboss did not interfere with R7 induction in the sensitized system. The inability of these two forms to inhibit R7 development is consistent with the observation that they are expressed at lower levels than EXboss (Fig. 2*c*).

In summary, we conclude that S2:boss cells or membranes prepared from these cells activate the sev tyrosine kinase. The extracellular domain of boss inhibits the boss-sev interaction both *in vivo* and *in vitro*. This is probably due to a low-affinity interaction between EXboss and the sev receptor. A soluble form of the membrane-bound, peptide/major histocompatibility complex binds to the T-cell receptor with low affinity¹⁸. Membrane-bound mutant forms of boss neither mediate adhesion to sev-expressing cells nor activate the sev tyrosine kinase. On the basis of these studies we propose that the seven-transmembrane domain region is essential for boss to function as a ligand for sev. The seven-transmembrane region may be required for the extracellular domain to assume a conformation that can bind with high affinity and activate sev. Alternatively, the sev receptor may bind to additional determinants located within the seven-transmembrane domain. We note that the G-protein coupled lutropin receptor, which also has a large extracellular domain and 7TM segments, contains binding sites in both these domains for large polypeptide hormones¹⁹. A third possibility is that the seven-transmembrane domain aids oligomerization which may be required for activity. Although sev has been reported to be a heterotetramer⁴, the oligomeric state of boss has not yet been determined. □

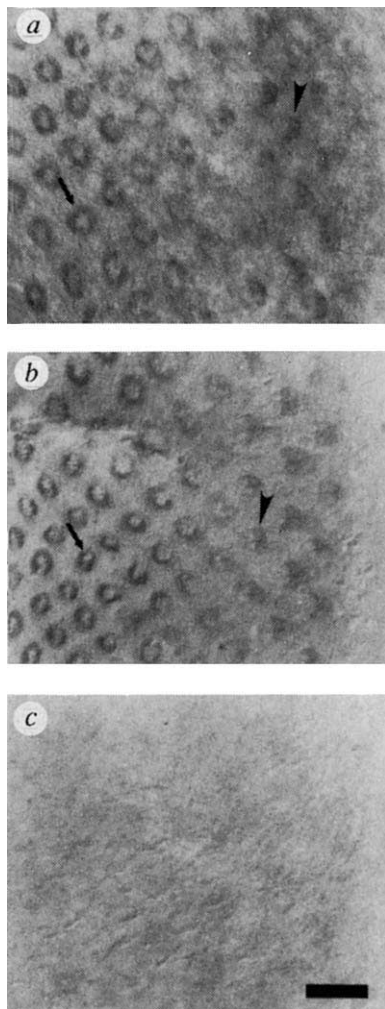


FIG. 4 Sev-dependent apical localization of mutant boss proteins. *a*, The apical localization of the EXboss protein, detected with pAbNT1, is remarkably similar to the apical localization of the sev receptor as illustrated in *b*, visualized with anti-sev mAb150C3 (ref. 15). *c*, The apical localization of EXboss is lost in a *sev*^{d2} mutant. Posterior is to the left. The morphogenetic furrow²¹ is at the right edge of each photograph. Arrowheads in *a* and *b* indicate staining in R3 and R4; arrows indicate cone cell staining. Scale bar, 12.5 μm .

METHODS. Immunolocalization of sev and mutant boss proteins in eye discs was done as previously described⁷. Third instar larvae containing the mutant protein constructs were incubated for 60 min at 37 °C, followed by a recovery period of 2 h. The heat shock and recovery were repeated and then the larvae were processed for eye disc staining.

Received 17 November; accepted 17 December 1992.

- Reinke, R. & Zipursky, S. L. *Cell* **55**, 321–330 (1988).
- Krämer, H., Cagan, R. C. & Zipursky, S. L. *Nature* **352**, 207–212 (1991).
- Hafen, E., Basler, K., Edstrom, J.-E. & Rubin, G. M. *Science* **236**, 55–63 (1987).
- Simon, M. A., Bowtell, D. L. & Rubin, G. M. *Proc. natn. Acad. Sci. U.S.A.* **86**, 8333–8337 (1989).
- Basler, K. & Hafen, E. *Cell* **54**, 299–311 (1988).
- Hart, A. C., Krämer, H., Van Vactor, D. L. Jr, Paidhungat, M. & Zipursky, S. L. *Genes Dev.* **4**, 1835–1847 (1990).
- Cagan, R. C., Krämer, H., Hart, A. C. & Zipursky, S. L. *Cell* **69**, 393–399 (1992).
- Massague, J. *J. biol. Chem.* **265**, 21393–21396 (1990).
- Ankelesaria, P. et al. *Proc. natn. Acad. Sci. U.S.A.* **87**, 3289–3293 (1990).
- Copeland, N. G. et al. *Cell* **63**, 175–183 (1990).
- Huang, E. et al. *Cell* **63**, 225–233 (1990).
- Zsebo, K. M. et al. *Cell* **63**, 213–224 (1990).
- Brannan, C. I. et al. *Proc. Natn. Acad. Sci. U.S.A.* **88**, 4671–4674 (1991).
- Planagan, J. G., Chan, D. C. & Leder, P. *Cell* **64**, 1025–1035 (1991).
- Banerjee, U., Renfranz, P. J., Hinton, D. R., Rabin, B. A. & Benzer, S. *Cell* **51**, 151–158 (1987).
- Tomlinson, A., Bowtell, D. L., Hafen, E. & Rubin, G. M. *Cell* **51**, 143–150 (1987).
- Van Vactor, D. L. Jr, Cagan, R. C., Krämer, H. & Zipursky, S. L. *Cell* **67**, 1145–1155 (1991).
- Matsui, K. et al. *Science* **245**, 1788–1791 (1991).
- Ji, I. H. & Ji, T. H. *J. biol. Chem.* **266**, 13076–13079 (1991).
- Basler, K. & Hafen, E. *Science* **67**, 931–934 (1989).
- Ready, D. F., Hansen, T. E. & Benzer, S. *Dev. Biol.* **53**, 217–240 (1976).

ACKNOWLEDGEMENTS. We thank the members of the Zipursky laboratory for discussions, M. Simon and G. Rubin for the SL2-sev cell line, and E. Hafen and K. Basler for the anti-sev antibodies and the plasmid pKB174 encoding sev^{met2242}. This research was supported by the Howard Hughes Medical Institute, a grant from USPHS NRSA (A.C.H.), an American Cancer Society Postdoctoral Fellowship (H.K.), a McKnight Foundation Development Award (S.L.Z.) and an NIH grant. S.L.Z. is an investigator of the Howard Hughes Medical Institute.

A yeast GTPase-activating protein that interacts specifically with a member of the Ypt/Rab family

Molly Strom, Petra Vollmer, Tjie J. Tan & Dieter Gollwitz*

Department of Molecular Genetics, Max-Planck Institute of Biophysical Chemistry, PO Box 2841, D-3400 Göttingen, Germany

MEMBERS of the Ras superfamily of GTP-binding proteins are involved in a variety of cellular processes, including signal transduction, cytoskeletal organization and protein transport^{1,2}. GTP-binding proteins of the Ypt/Rab family direct vesicular protein transport in the secretory and endocytic pathways in the yeast *Saccharomyces cerevisiae* (Ypt proteins) and in mammalian systems (Rab proteins)^{3,4}. The cellular activity of monomeric GTP-binding proteins is influenced by proteins that regulate GDP/GTP exchange and GTP hydrolysis⁵. GTPase-activating proteins (GAPs) can increase the slow intrinsic GTPase activity of GTP-binding proteins by several orders of magnitude^{6,7}. As GAPs modulate the activity of GTP-binding proteins, they are thought to give a biochemical handle on the functioning of Ypt/Rab proteins in transport vesicle budding and docking or fusion at donor and acceptor membranes^{1,2}. We report here the first cloned GTPase-activating protein for the Ypt/Rab protein family. The gene, *GYP6* (GAP of Ypt6 protein), encodes a protein of 458 amino acids which is highly specific for the Ypt6 protein and shows little or no cross-reactivity with other Ypt/Rab family members or with H-Ras p21.

The recently identified Ypt6 protein (Ypt6p) of *S. cerevisiae*, a homologue of the fission yeast Ryh1p (ref. 8) and the Golgi-located mammalian Rab6p (ref. 9), is not essential for cell viability and seems to be involved in vacuolar protein sorting or transport (L. Hengst, H. Wichmann and D. G., unpublished results). We isolated the gene encoding the GAP activity for Ypt6p by taking advantage of the *S. cerevisiae* genome's relatively small size and the fact that several enzyme-encoding yeast genes have been cloned based on overexpression when present on a multi-copy plasmid¹⁰. We screened individual protein extracts prepared from ~2,500 single *S. cerevisiae* transformants for increased GAP activity using [γ -³²P]GTP-loaded Ypt6p as

TABLE 1 Substrate specificity of Gyp6p

Protein	Effector region	Fold activation of intrinsic GTPase
Ypt6	-YQATIGIDF-	>100
Ypt7	-YKATIGADF-	11
Ypt1	-YISTIGVDF-	—
Sec4	-FITITIGIDF-	—
Ypt3A	-SKSTIGVEF-	—
H-Ras	-YDPTIEDSY-	—
Rab2	-HDLTIGVEF-	—

The various GTP-binding proteins expressed and purified from *E. coli*³⁵, (except H-Ras p21, which was a gift from A. Wittinghofer) were loaded with [γ -³²P]GTP (ref. 32). Equivalent amounts of [γ -³²P]GTP-bound protein, based on c.p.m. incorporated, were incubated with 3 different concentrations (0.13 μ g μ l⁻¹, 0.4 μ g μ l⁻¹, and 1.3 μ g μ l⁻¹) of a partially purified Gyp6 protein, or with buffer only (50 mM Tris-HCl, pH 8.0, 1 mM DTT, 5.0 mM MgCl₂). 2 × 10 μ l samples were taken at 0, 5, 10, 15 and 20 min for each protein, except for Ypt6p, for which 0, 1, 2, 5, 10 and 20 min samples were taken, and trapped onto BA 85 filters. Filters were washed and counted for ³²P. The acceleration of the intrinsic GTPase activity was measured from the linear range of the curve and compared with the intrinsic GTPase activity. GTPase activity measurements were made at least twice with each substrate protein. The partially purified Gyp6p fraction was prepared by passage of *E. coli* protein (prepared from 8 l of PET11a-GYP6-transformed cells) over DEAE-Sephacel (Pharmacia), eluted with a 0–500 mM NaCl gradient in 10 mM Tris-HCl, pH 8.0, 2 mM MgCl₂, 1 mM DTT, 1 mM EGTA, and the active fractions pooled (140 ml) and concentrated to ~20 ml by Amicon YM10 membrane filtration (Amicon). This sample was then passed over a Superdex-200 (Pharmacia) column in 10 mM Tris-HCl 8.0, 1 mM EGTA, 1 mM DTT, 2 mM MgCl₂, 0.6 KCl. Ten-ml fractions were collected and active fractions individually concentrated to 1 ml by Amicon YM Membrane filtration.

a substrate. A single colony was isolated that had ~10- to 20-fold more Ypt6p-GAP activity than wild-type cells (Fig. 1). We subcloned smaller fragments of the original 8-kilobase (kb) insert contained in the recombinant plasmid, and isolated the *GYP6* gene on a 3.3-kb *Clal*-*XbaI* fragment. As seen in Fig. 1a, protein extracts prepared from yeast cells transformed with the vector pRS326 containing the 3.3-kb genomic DNA fragment (pRS326-GYP6), but not with the insert-free vector, strongly activated the Ypt6p intrinsic GTPase activity (for the same degree of activation of the intrinsic Ypt6p GTPase activity as in wild-type yeast protein extracts, about 50-fold less protein was required). Analysis of the protein-bound nucleotide by thin-layer chromatography showed that this was induced due to a GAP-like activity as the [α -³²P]GTP bound to Ypt6p after incubation with extracts from pRS326-GYP6-transformed cells was hydrolysed specifically to GDP (Fig. 1c). To prove unequivocally that the DNA fragment we isolated encoded the GAP, we expressed the *GYP6* gene product in *Escherichia coli* and tested for Ypt6-GAP activity (Fig. 1b and c). The open reading frame deduced from the DNA sequence (Fig. 2) represents Ypt6-GAP, because in incubations with *E. coli* extracts, as with *S. cerevisiae* protein extracts, the GTP bound to Ypt6p was efficiently hydrolysed to GDP (Fig. 1c). This proved that Gyp6p is a GTPase-activating protein and not a GDP/GTP exchange factor, or a nonspecific protease or phosphatase. The Gyp6 protein encoded on the 3.3-kb DNA fragment was established to lie within a 2.4-kb *PmlI*-*XbaI* DNA fragment. This was sequenced and is shown in Fig. 2 with the predicted amino-acid sequence of the 458-amino-acid protein. Comparison of the amino-acid sequence using the Fasta program³⁷ revealed no significant homology to known sequences, including the GAPs for Ras^{11–15}, Rap¹⁶, Rac¹⁷, and Rho¹⁷ proteins.

To assess the functional significance of Gyp6p, the gene encoding the GTPase-activating protein was disrupted by insertional mutagenesis (Fig. 2b). Unlike *ypt6* deletion mutants which exhibit temperature-sensitive growth and fragmented vacuoles (L. Hengst, H. Wichmann and D. G., unpublished results), *gyp6* disruptions grew at a normal rate and without apparent

* To whom correspondence should be addressed.

phenotypic alterations. Likewise, overexpression of Gyp6p from a multicopy plasmid did not result in an easily scoreable phenotype. A *ypt6/gyp6* double-deletion mutant behaved like *ypt6* null mutants. Less severe consequences resulting from the lack of function of a GTPase-activating protein, as opposed to the functional loss of the GTP-binding protein itself, is also known from the Ras system in *S. cerevisiae*. Cells with the simultaneous disruption of the *IRA1* and *IRA2* genes encoding the Ras-GAP activities retain viability¹⁵, whereas the deletion of *RAS1* and *RAS2* leads to lethality^{18,19}.

As the Ypt/Rab family consists of a large number of structurally and functionally distinct proteins, the specificity of Ypt/Rab protein GAPs is one of the major questions in relationship to their possible cellular functions: are they important only as regulators of GTP hydrolysis or do they also have other roles, possibly in donor/acceptor membrane recognition or membrane fusion processes? It is conceivable that one GAP could regulate the GTPase activity of more than one GTP-binding protein, as was recently shown with the Rho-GAP of the Ras-GAP-associated p190, which accelerates the intrinsic GTPase activity of five Rho family members²⁰. Perhaps, in the case of the Ypt/Rab proteins, there exist organelle- or pathway-specific GAPs. To test for substrate specificity, we used *E. coli*-expressed Gyp6p to avoid having to account for the presence in yeast extracts of other endogenous GAP activities for the various proteins analysed. As shown in Table 1, we tested all known Ypt proteins of *S. cerevisiae*, Ypt1p (ref. 21), Sec4p (ref. 22), Ypt7p (ref. 23), Ypt6p and Ypt3Ap (our unpublished results). The mammalian H-Ras p21 (ref. 24) and Rab2p (ref. 25) were included in this analysis. Ypt7p was the only other of the proteins tested whose

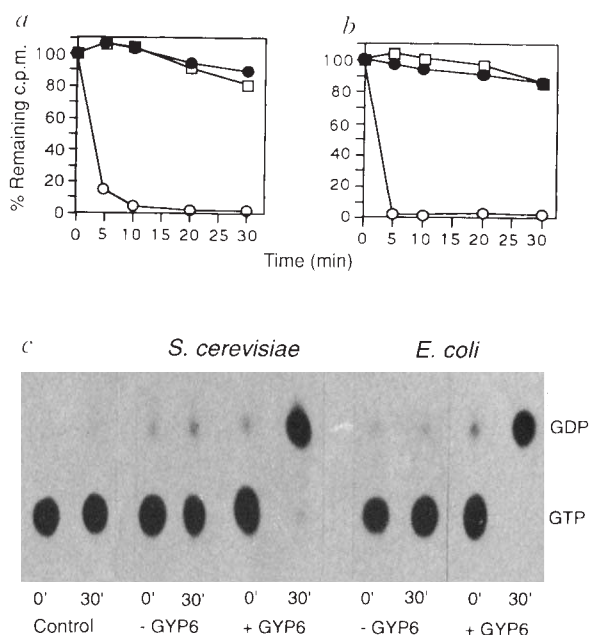
slow intrinsic GTPase activity was accelerated by Gyp6p, but this increase was at least an order of magnitude lower than that seen with Ypt6p. We also tested the activity of the Gyp6 protein overproduced in *S. cerevisiae* with the proteins Ypt1p, Ypt3Ap, Sec4p, Ypt6p and Ypt7p and obtained the same results (data not shown), showing that the *E. coli*-produced protein is not lacking modifications that would impair the *in vitro* analysis of GAP activity.

Table 1 also lists the effector regions for the GTP-binding proteins tested. Not only is the effector region involved in GAP interaction^{7,26-28}, it is also used as an indicator for functional homology between GTP-binding proteins of different organisms^{8,29-31}. Based on sequence differences of the effector regions, it was not surprising that Gyp6p had no effect on the GTPase activities of Ypt1p, Ypt3Ap, Sec4p, or the mammalian Rab2 and Ras proteins. Although the effector region of Ypt7p is most similar to that of Ypt6p (two amino-acid differences), the acceleration of the Ypt7p GTPase activity by Gyp6p was unexpected because these two proteins are functionally distinct²³. Although we cannot eliminate the possibility that Gyp6 *in vivo* acts as GAP for both proteins, preliminary fractionation of soluble yeast proteins by ion-exchange chromatography reveals a Ypt7p-GAP activity which is separable from Gyp6p activity; the latter is undetectable in the *ypt6* deletion mutant (data not shown).

The high degree of specificity of Gyp6p, and recent demonstrations of GAP activities for *S. cerevisiae* Ypt1p (ref. 32) and Sec4p (ref. 33), indicate that there are several different GAPs, if not a specific GAP, for each Ypt protein. This also has interesting implications for their cellular function. As there

FIG. 1 Expression of Gyp6p activity in *S. cerevisiae* and *E. coli*. **a**, [γ -³²P]GTP-loaded Ypt6p was incubated at 30°C with buffer only (●), or protein extracts (0.3 μ g μ l⁻¹) prepared from the yeast strain BJ3505 containing plasmids pRS326-GYP6 (○) or pRS326 without insert (□). 2 $\times$ 10 μ l samples were filtered through nitrocellulose (BA 85; Schleicher & Schuell) at 0, 5, 10, 20 and 30 min time points and filters were washed and counted for ³²P. Values represent the fractional per cent of the counts at *t* = 0 (~350,000 c.p.m.). **b**, [γ -³²P]GTP-loaded Ypt6p was incubated with protein extracts (0.3 μ g μ l⁻¹) prepared from the *E. coli* strain BL21 containing plasmids PET11a-GYP6 (○), PET11a (□), or with lysis buffer (●). Samples were processed as in **a**. **c**, Protein extracts from GYP6-transformed yeast (0.1 μ g μ l⁻¹) or *E. coli* (0.06 μ g μ l⁻¹) and a buffer control, were incubated at 30°C with [α -³²P]GTP-loaded Ypt6p. A 30- μ l sample was taken at 0 and 30 min, washed, and the bound guanine nucleotides were then dissociated from protein and separated by thin-layer chromatography. Control, buffer alone; -GYP6 is protein prepared from insert-free vector-transformed strains (pRS326 in yeast and PET11a in *E. coli*); +GYP6 is protein prepared from strains transformed with vectors containing the GYP6 gene (pRS326-GYP6 in yeast and PET11a-GYP6 in *E. coli*).

METHODS. Library screening: Yeast cells (strain BJ3505: *a pep4::HIS3 prb1- Δ 1.6R lys2-208 trp1- Δ 101 ura3-52 gal2 can1*) were transformed with DNA from a yeast genomic library³⁴. Single colonies were picked and grown in 5 ml uracil-free liquid medium (2% glucose, 0.67% yeast nitrogen base without amino acids, 0.5% peptone 140) for 20–24 h at 30°C. Washed cells were resuspended in 120 μ l lysis buffer (10 mM Tris-HCl, pH 8.0, 2 mM MgCl₂, 1 mM EGTA, 1 mM DTT, 1 mM PMSF), and shaken with 1 vol glass beads for 1 h at 4°C on an Eppendorf mixer. The 8,000g supernatant of disrupted cells was used for GAP assays. GAP assays and [γ -³²P]GTP loading of Ypt6p were as described³², except 42 nM [γ -³²P]GTP (6,000 Ci mmol⁻¹; Amersham) was used and the [γ -³²P]GTP-labelled protein was diluted sixfold in buffer A (50 mM Tris-HCl, pH 8.0, 2 mM EDTA, 1 mM DTT, 5 mM MgCl₂, 0.5 mg ml⁻¹ BSA). 50 μ l reactions (10 μ l diluted [γ -³²P]GTP-labelled Ypt6p, 10 μ l protein extract, 1 mM each of ATP and GTP and buffer A to 50 μ l) were incubated for 30 min at 30°C in 96-well microtitre plates. 2 $\times$ 10 μ l samples were filtered onto nitrocellulose using a Biorad 96-well dot-blot apparatus and washed. The nitrocellulose filters were dried, autoradiographed, and analysed by eye for reduction in intensity of the radioactive spots. Proteins: Ypt6p was expressed in pLN vector in *E. coli* and purified as described³⁵. Yeast protein extracts were prepared as in library screening, except that the extracts were vortexed for 6 $\times$ 1 min at 3-min intervals at 4°C. For preparing *E. coli* protein extracts, 40 ml of cells were grown to an absorbance of 0.3 at 600 nm in LB medium, and isopropyl- β -D-thio-



galactoside (IPTG) was added to 1 mM. After 5 h induction, cells were pelleted and resuspended in 400 μ l of lysis buffer (50 mM Tris-HCl 7.5, 1 mM EDTA, 1 mM DTT, 0.2 mM PMSF, 1% aprotinin, 0.2 mg ml⁻¹ lysozyme) and incubated on ice for 30 min, then DNase I (10 μ g ml⁻¹) and 10 mM MgCl₂ were added and incubation continued for 20 min on ice, followed by centrifugation at 10,000g for 5 min. Plasmids: For yeast experiments, a 3.3-kb DNA fragment harbouring GYP6 was subcloned into the vector pRS326. pRS326 was constructed from pRS306³⁶, which contains the *URA3* gene as a yeast selectable marker, by inserting the *EcoRI* 2- μ DNA-containing fragment from YEp24 into the *AatI* site of pRS306. Genomic fragments were cloned into the unique *SmaI* site of pRS326. The PET11a-GYP6 vector was created by introducing a *NdeI* site at the ATG initiation codon of GYP6 and cloning into the *NdeI* site of PET11a (Novagen).

a

```

-450      CTGCGGTTTTCTGAACACTCACCTTTAGIATTCCTTTTTCTTTTTT
-400  TTCTATTTTTTTTCGACATAATTATTAATTTTTCTGAATTAATTTACTGACGCAAGTCTGACACAGAAAAATTC
-320  CGTCTTATTTTGGCCATGCCAGCGTAGAATACGCACAGGTGACGCATGCTCCAAGTGTTATTTGTAACGGAAGGTCG
-240  CTGGTTTAGGCGCTAAGCGGTGCTAAGGAAACCGGATATATGTTACCGCTTCTCCTACGCGATTACTAGTTCCTT
-160  TTTTCTTTTTTTTCAATTTTCCATTGAGTTTCCATAATCGGTACAGCAGCGTTTAATTTTATTACCTTTTTCAGAAC
-80   ATGGGGGAGAGTTGTCAAGAGAAATGGCATAACATAGAGAGGTTAGCTGTGCGGTTGTATTTTTCGTTATGCAACC

1   ATG AAG GAT GTC TTA CAA TGG GCG ATT GAC CAC TAC GAG ACA AGA GAC CAA CTG GTG CAG
1   Met lys asp val leu gln trp ala ile asp his tyr glu thr arg asp gln leu val gln

61  AAG GGA ATT TGS AAA GGT GAA CTC TAC CAT GAC TCT ATC TTT AAA GAA AAT AGT AGA GGC
21  lys gly ile trp lys gly glu thr tyr his asp ser ile phe lys glu asn ser arg gly

121 TGG CTC TGS AAA GCT TTG CTT CTG TGT GAT GAA AAC AAC AAC TGC TTG TTA ACT GAT AAC
41  trp leu trp lys ala leu leu leu cys asp glu asn asn asn cys leu leu thr asp asn

181 TTT AAA GGT TTG GAC CTT AAC CAG TTC GGC TTG GTG CCA GTG CCT ATG CTG GCA GAT GGC
61  phe lys gly leu asp leu asn gln phe gly leu val pro val pro met leu ala asp gly

241 GAC AAC TAT GAC GAA AAT CAT AAT GCA AAT GTA CCT AAA AGG GTC TTG CAC TCG AAC GTT
81  asp asn tyr asp glu asn his asn ala asn val pro lys arg val leu his ser asn val

301 TCC TCC TCA GTG GGA ATC AGG AGG CTG ACT CCC GTC GAG GCT GTT GAA AAA CAT CCC TTG
101 ser ser ser val gly ile arg arg leu thr pro val glu ala val glu lys his pro leu

361 TCT GAT GAC AAT GAC AAA ACT AAA GGT TCT TTG TCC AAG GGA AGT GAT GAA AAG CCG TTG
121 ser asp asp asn asp lys thr lys gly ser leu ser lys gly ser asp glu lys pro leu

421 ACA CTG AGG GAA ACT CTG GAA ATT ATA GAT TTG GAC CTC TCA AGG ATA ATG CTT GAC GAT
141 thr leu arg glu thr leu glu ile ile asp leu asp leu ser arg ile met leu asp asp

481 ATA TTT CAG GAA CCC AAG GTT CAT GCT CAA ATG AGA CAA CTT TTG TAC AAT TAC CTG CTT
161 ile phe gln glu pro lys val his ala gln met arg gln leu leu tyr asn tyr leu leu

541 ATA CAT CAA AGC GAA CAC CTA CAG TAT AAA CAA GGG TGC CAC GAA ATA TTA AGC GTC ATT
181 ile his gln ser glu his leu gln tyr lys gln gly phe his glu ile leu ser val ile

601 TAT CTG CAA TTA TAT CAC GGT ACA GAT CTC GAT AAT ACC GAT TTG CAA AAC GTT CTG ATT
201 tyr leu gln leu tyr his gly thr asp leu asp asn thr asp leu gln asn val leu ile

661 ATC TTT AAC AAG CTA ATG AAC CAA ATC GAA CCG ATT TTT TAT AAC GAA GAA AAT TTG ATT
221 ile phe asn lys leu met asn gln ile glu pro ile phe tyr asn glu glu asn leu ile

721 AAC TGG GAT AAG AGA GTA TTC ACG AAA ATC TTT AGG ATT TGT TTG CCA GAC TTG TTT TCA
241 asn trp asp lys arg val phe thr lys ile phe arg ile cys leu pro asp leu phe ser

781 AAG GTT TTT TAC CAA CCA CCA AAA ACG GGC AGC GGG AAG AAA AAA AAC GTA GAC CAT CTG
261 lys val phe tyr gln pro pro lys thr gly ser gly lys lys lys asn val asp his leu

841 ATT CAC TCC AAC CTA ATC TGG CTC ATC CGA TGG ACA AGG TTG TTA TTT TTA AGA GAA TTA
281 ile his ser asn leu ile trp leu ile arg trp thr arg leu leu phe leu arg glu leu

901 CCC TTA AAA TAT GTG CTA ATC CTT TGG GAC CAC GTC CTG ACG TTC AAT TAT CCC TTG GAT
301 pro leu lys tyr val leu ile val trp asp his val leu thr phe asn tyr pro leu asp

961 ATT TTC ATT GCA TGC ACT ATC ATC ACC CTT TTG TTA TCC ATT TAC GAT GAG CTT CAC GAA
321 ile phe ile ala cys thr ile ile thr leu leu leu ser ile tyr asp glu leu his glu

1021 CTA GTG AGC CAA GGT GAC TAT GAA CAT ACG AAT AAT AAC GAT GAA TTT GTT GAA TTG ATT
341 leu val ser gln gly asp tyr glu his thr asn asn asn asp glu phe val glu leu ile

1081 TTG CAC TTC AAG AAA ATT TTT GAA AAG GAA GAT GCC AGT AAA GAT GAT GAA AAG TTT TTA
361 leu his phe lys lys ile phe glu lys glu asp ala ser lys asp asp glu lys phe leu

1141 GAC CTG TGT AAG GTC ACT GGA AAT TTA TGT GAA TTG TGG TAT GGA AAA AAT TAC GAC GAT
381 asp leu cys lys val thr gly asn leu cys glu leu trp tyr gly lys asn tyr asp asp

1201 ATG AGA TTG ATT TGC GAC ACT TTT ATA AAT GCG AAG TTA GGC ATA AAG ACT ACT GAT GTT
401 met arg leu ile cys asp thr phe ile asn ala lys leu gly ile lys thr ser asp val

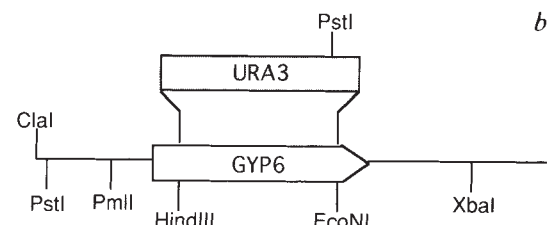
1261 TTA AGT ATG GAA ACA GCA AAG TTG ACT ATT GAT CCT AAT AGA CAA TCA CTC GAA AAC AAA
421 leu ser met glu thr ala lys leu thr ile asp pro asn arg gln ser leu glu asn lys

1321 CTG AGG GAG AGG GTA AGG CAA ACA ATT TTG AAG AAC AAA AAG AAA ATA AGC CAA TAA
441 leu arg glu arg val arg gln thr ile leu lys asn lys lys lys ile ser gln OCH

1382 TTCTTGTCATTTACCGTGTGTATTTGTTTGAATCCCGAGCGTCACACCCGACGACAAAAATCAAAAAAAAAAAAAA
1462 AAACCCAGCTACCACCAACAACAGATAGCCACATGTCTATTTCTCTAATGAGCTGGTTTCCCTCTGATATTTTGAT
1542 ATTGCAGCTCGAGCTATATTTGACTTCAATGTCTTGGCATAAAATATATAGCATAATAGATAAGGTGIAAGTTGACAGT
1622 AAACGAGTACTATTGACCAAGAAATCAATGAGAGAGATAGAGAGAAAAAATAGAAAAAAGGGGAGATCATTTTGGCAGAT
1702 CAATAAATAAGTTTGGCCCTCCAAAAATAAATAAATAAATAAAGAAATTAATACGATCTGACAGCTGGAGGGGACTGGGGC
1782 GCAATCTCATCTAAAGIATGTGGGATGACATTCCTGTACTTGATTTTACAGGTGTTGATGTCCCTTTTGGCATGATAA
1862 TGTTGTCTTTCTCCAAATTCACATCATCTCTTTTGATAGTCTCTCTCGCATGAGCACCAGAGACTCCTTTCTTTGG
1942 AAGCAGAAACAGCCGTTTGTGTGGCAAGTAAGTAAATTTTGAATTCAGAGTTTCTACGAGATCAGAAATTCACAGAT
2022 CATTCACCTTATCAGAAACGTGACATCTCTCAAAATGCTTGTCCACTTCAGTAATATTTCTAACACCTTCA

```

FIG. 2 a, Nucleotide and predicted amino-acid sequence of *GYP6*. DNA from the *Clal*-*XbaI* 3.3-kb fragment containing the *GYP6* gene was sequenced. Based on the identification of an open reading frame in this 3.3-kb fragment, a smaller 2.4-kb *PmlI*-*XbaI* fragment was subcloned into pRS326 and extracts from yeast transformed with this fragment also showed increased Gyp6p activity. The DNA sequence, and predicted amino-acid sequence, from this 2.4-kb fragment are shown here. EMBL accession number, X68506. b, Disruption of *GYP6*. The *HindIII*-*EcoNI* fragment (removing amino acids 45 to 384) was replaced with the *SmaI*-*Clal* fragment of the yeast vector YEp24 containing the *URA3* selectable marker. Disruption of *GYP6* was confirmed by Southern blot analysis.



are different GTP-binding proteins that associate with different membrane-enclosed structures (ER, Golgi, endosomes, vesicular carriers), it is conceivable that the GAPs could play a role in membrane recognition or fusion processes. Genetic and biochemical studies in progress should elucidate the role of Gyp6p in protein transport or sorting and the mechanism for its interaction with Ypt6p. □

Received 1 October; accepted 30 December 1992.

1. Bourne, H. R., Sanders, D. A. & McCormick, F. *Nature* **348**, 125–132 (1990).
2. Hall, A. *Science* **249**, 635–640 (1990).
3. Balch, W. *Trends Biochem. Sci.* **15**, 473–477 (1990).
4. Pfeffer, S. R. *Trends Cell Biol.* **2**, 41–46 (1992).
5. Bourne, H. R., Sanders, D. A. & McCormick, F. *Nature* **349**, 117–127 (1991).
6. Trahey, M. & McCormick, F. *Science* **238**, 542–545 (1987).
7. Frech, M. *et al. Science* **249**, 169–171 (1990).
8. Hengst, L., Lehmeier, T. & Gallwitz, D. *EMBO J.* **9**, 1949–1955 (1990).
9. Goud, B., Zahraoui, A., Tavitian, A. & Saraste, J. *Nature* **345**, 553–556 (1990).
10. Rine, J. *Meth. Enzym.* **194**, 239–251 (1991).
11. Vogel, U. S. *et al. Nature* **335**, 90–93 (1988).
12. Trahey, M. *et al. Science* **242**, 1697–1700 (1988).
13. Xu, G. *et al. Cell* **63**, 835–841 (1990).
14. Martin, G. A. *et al. Cell* **63**, 843–849 (1990).

15. Tanaka, K. *et al. Molec. cell. Biol.* **10**, 4303–4313 (1990).
16. Rubinfeld, B. *et al. Cell* **65**, 1–20 (1991).
17. Diekmann, D. *et al. Nature* **351**, 400–402 (1991).
18. Tatchell, K., Chaleff, D. T., DeFeo-Jones, D. & Scolnick, E. M. *Nature* **309**, 523–527 (1984).
19. Kataoka, T. *et al. Cell* **37**, 437–445 (1984).
20. Settleman, J., Albright, C. F., Foster, L. C. & Weinberg, R. A. *Nature* **359**, 153–154 (1992).
21. Gallwitz, D., Donath, C. & Sander, C. *Nature* **308**, 704–707 (1983).
22. Salminen, A. & Novick, P. *Cell* **49**, 527–538 (1987).
23. Wichmann, H., Hengst, L. & Gallwitz, D. *Cell* **71**, 1131–1142 (1992).
24. Capon, D. J., Chen, E. Y., Levinson, A. D., Seeburg, P. H. & Goeddel, D. V. *Nature* **302**, 33–37 (1983).
25. Touchot, N., Chardin, P. & Tavitian, A. *Proc. natn. Acad. Sci. U.S.A.* **84**, 8210–8214 (1987).
26. Calés, C., Hancock, J. F., Marshall, C. J. & Hall, A. *Nature* **332**, 548–551 (1988).
27. McCormick, F. *Cell* **56**, 5–8 (1989).
28. Becker, J., Tan, T. J., Trepte, H.-H. & Gallwitz, D. *EMBO J.* **10**, 785–792 (1991).
29. Haubruck, H., Prange, R., Vorgias, C. & Gallwitz, D. *EMBO J.* **8**, 1427–1432 (1989).
30. Miyake, S. & Yamamoto, M. *EMBO J.* **9**, 1417–1422 (1990).
31. Haubruck, H., Engelke, U., Mertins, P. & Gallwitz, D. *EMBO J.* **9**, 1957–1962 (1990).
32. Tan, T. J., Vollmer, P. & Gallwitz, D. *FEBS Lett.* **291**, 322–326 (1991).
33. Walworth, N. C., Brennwald, P., Kabcenell, A. K., Garrett, M. & Novick, P. *Molec. cell. Biol.* **12**, 2017–2028 (1992).
34. Carlson, M. & Botstein, D. *Cell* **28**, 145–154 (1982).
35. Wagner, P. *et al. EMBO J.* **6**, 2373–2379 (1987).
36. Sikorski, R. S. & Hieter, P. *Genetics* **122**, 19–27 (1989).
37. Pearson, W. R. & Lipman, D. J. *Proc. natn. Acad. Sci. U.S.A.* **85**, 2444–2448 (1988).

ACKNOWLEDGEMENTS. We thank L. Hengst, M. Benli, F. Döring, W. Laufer and H. Wichmann for purified Ypt and Rab proteins, U. Welscher-Altschäffel and P. Mienkus for technical help, and A. Wittinghofer for H-Ras p21. M.S. acknowledges the receipt of a HFSP fellowship. This work was supported in part by the Deutsche Forschungsgemeinschaft.

Conversion of lytic to persistent alphavirus infection by the *bcl-2* cellular oncogene

Beth Levine^{*†}, Qi Huang[‡], John T. Isaacs^{§||},
John C. Reed[¶], Diane E. Griffin^{*#}
& J. Marie Hardwick^{‡##}

Departments of * Medicine, ‡ Pharmacology and Molecular Sciences, § Oncology, || Urology, and # Neurology, Johns Hopkins University School of Medicine, Baltimore, Maryland 21287-7681, USA
¶ La Jolla Cancer Research Foundation, Cancer Research Institute, La Jolla, California 92037, USA

† Present address: Department of Medicine, Columbia University College of Physicians and Surgeons, New York, New York 10032, USA

** To whom correspondence should be addressed

LITTLE is known about virus–host cell interactions that regulate the lytic potential of viruses during productive replication. Sindbis virus (SV), a single-stranded positive-sense RNA virus in the alphavirus genus (family *Togaviridae*), results in lytic infection in most vertebrate cell lines, but persistent productive infection in post-mitotic neurons¹. The cellular oncogene *bcl-2*, which encodes an inner mitochondrial membrane protein of *M_r* 26,000 (ref. 2), blocks programmed cell death (apoptosis) in neurons³. We therefore investigated whether SV infection induces programmed cell death in non-neuronal cells, and if so, whether virus-induced programmed cell death can be blocked by transfection with *bcl-2*. We demonstrate that SV infection of baby hamster kidney (BHK-2), mouse neuroblastoma (N18), and rat prostatic adenocarcinoma (AT-3) cells results in programmed cell death, whereas SV infection of *bcl-2*-transfected AT-3 cells results in long-term persistent productive infection. Thus cellular *bcl-2* oncogene expression plays a role in the establishment of persistent viral infection by blocking virus-induced programmed cell death.

Classically, lytic viral replication is believed to induce cellular necrosis either by destruction of the cell membrane, cell fusion, or lethal parasitism of host cell transcriptional or translational machinery. Cytotoxicity due to acute human immunodeficiency virus (HIV) infection^{4,5} and adenovirus E1A expression⁶ is associated with apoptosis, an active cell-suicide process characterized by distinct morphological features and endonuclease-induced DNA fragmentation into multimers of 180–200 base pairs (bp)^{7,8}. We studied SV-induced cell death in mammalian cells. The three cell lines BHK-21, N18 and AT-3 were infected

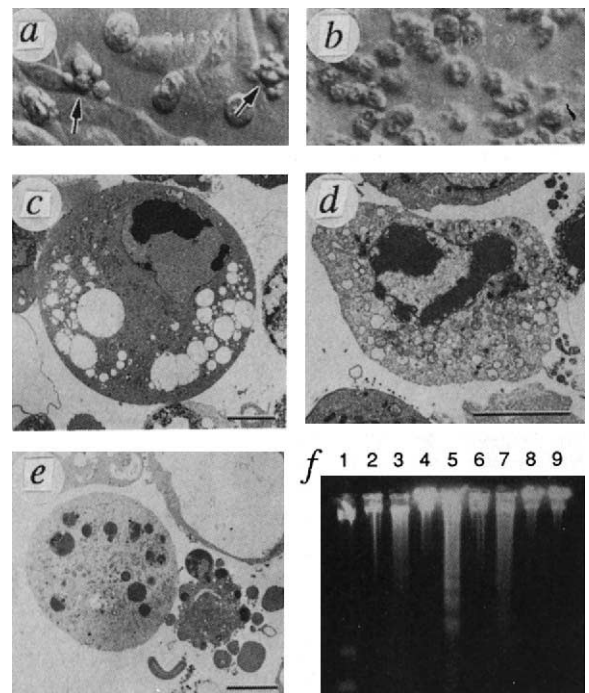
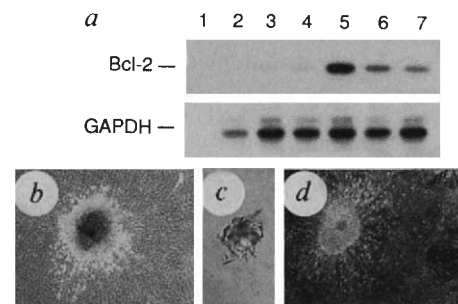


FIG. 1 Induction of apoptosis by SV infection in BHK, AT-3 and N18 cells. *a* and *b*, Videomicrographs²¹ of N18 cells. 10^6 cells grown in 10% fetal bovine serum were transferred to medium containing 2% fetal bovine serum and infected with wild-type SV strain AR339 (multiplicity of infection (MOI) of 1 plaque-forming unit per cell) at time 0, and observed by videomicroscopy for 72 h. *a*, A representative frame taken 24 h after infection shows typical apoptotic cells (arrows) with nuclear condensation, membrane blebbing and cytoplasmic condensation; in *b*, cells are dead and fragmented 48 h after infection. Similar results were obtained for SV-infected BHK-21 at AT-3 cells (data not shown). *c–e*, Electron micrographs of BHK-21 (*c*), AT-3 (*d*), and N18 (*e*) cells 24 h after SV infection (MOI=1) showing apoptotic morphology (see text for details). Scale bars, 5 μ m. *f*, Nucleosomal laddering in SV-infected BHK-21, AT-3 and N18 cells. Between 5×10^5 (BHK-21) and 10^6 (N18, AT-3) cells were collected 18 h after SV infection and 1.5×10^5 neuron-enriched 6-week-old rat dorsal root ganglion (DRG) cells were collected 48 h after SV infection. DNA laddering was detected by *in situ* extraction and agarose gel electrophoresis as described²². Both viable adherent and dying cells with reduced adherence were harvested to maintain total cell numbers constant. Lane 1, 123-bp ladder standard; lanes 2 and 3, mock and SV-infected BHK-21 cells; lanes 4 and 5, mock and SV-infected AT-3 cells; lanes 6 and 7, mock and SV-infected N18 cells; lanes 8 and 9, mock and SV-infected rat DRG cells.

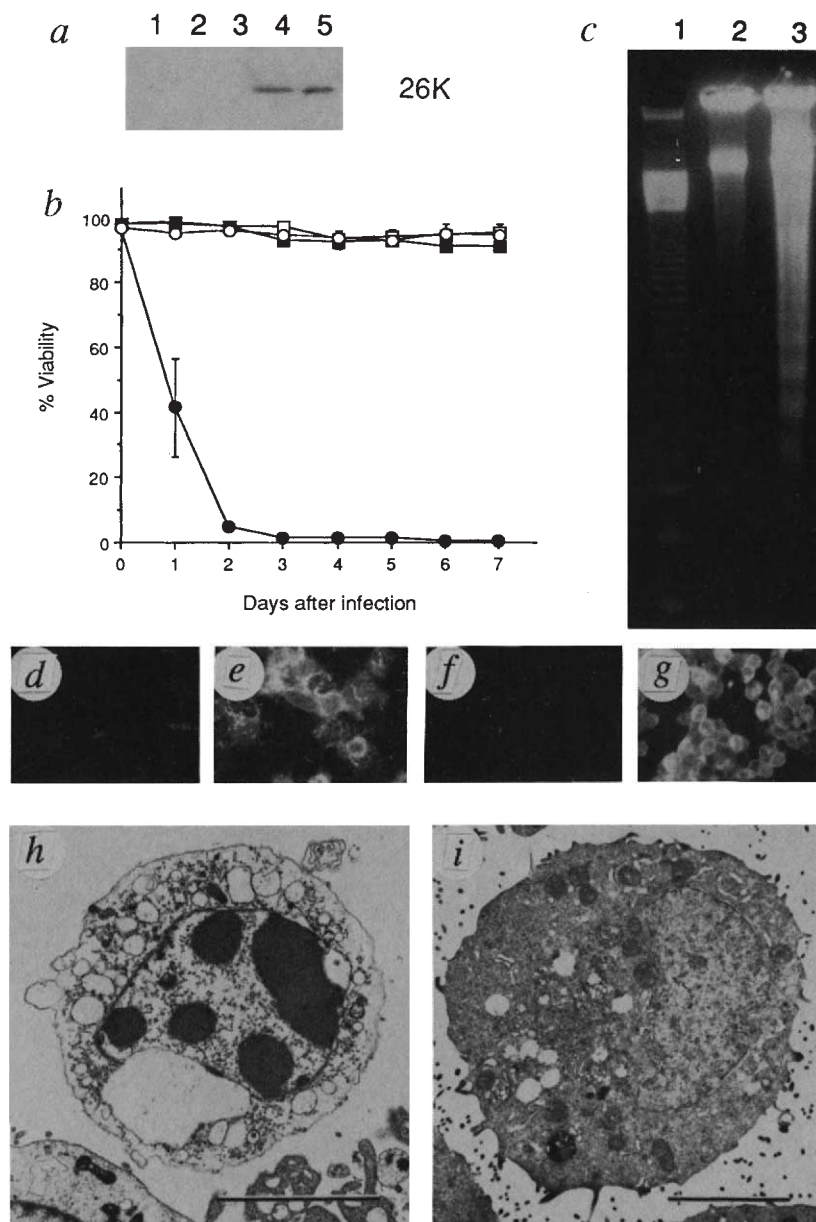
FIG. 2 Relationship between levels of *bcl-2* mRNA expression and susceptibility to SV lytic replication in cultured neurons. *a*, Southern blot analysis of *bcl-2* amplified by reverse transcription and polymerase chain reaction (RT-PCR) and glyceraldehyde phosphate dehydrogenase (GAPDH) mRNA from freshly explanted rat DRG cultures (lanes 2–4) and differentiated 6-week-old neuron-enriched DRG culture (lanes 5–7). Lane 1 represents PCR control. *b*, Photomicrograph of control uninfected 5-day-old rat DRG culture. *c*, Photomicrograph taken 2 days after SV infection (MOI=0.1) of 3-day-old DRG culture revealing dead neurons. *d*, Photomicrograph taken 7 days after SV infection (MOI=0.1) of 6-week-old DRG culture revealing neurons without detectable cytopathic effects.

METHODS. DRG cultures were prepared as described¹. Total RNA was extracted from 40 DRG (~30,000 cells) using the RNAzol method (TM Cinna Scientific, Friendswood, Texas). cDNA was synthesized using oligo(dT) priming and avian myeloblastosis virus reverse transcriptase at 94 °C for 30 s, annealing at 60 °C for 60 s, and extension at 72 °C for 50 s. PCR products were transferred to nylon membranes and Southern blots were hybridized with ³²P-labelled internal oligonucleotide probes. Oligonucleotide sequences for rat *bcl-2* PCR primers and for the internal oligonucleotide probe used in Southern blotting corresponded to areas of 100 homology between



published mouse¹² and human²³ *bcl-2* sequences; plus-strand primer (mouse 1,895–1,914), 5'-AGAGGGGCTACGAGTGGGAT-3'; minus-strand primer (mouse 2,330–2,349), 5'-CTCAGTCATCCACAGGGCGA-3'; internal oligonucleotide probe (mouse 2,245–2,274), 5'-ACCGAACTCAAAGAA-CAATCCTCCC-3'. GAPDH nucleotide sequences were as described²⁴.

FIG. 3 Effect of *bcl-2* transfection on susceptibility of AT-3 cells to SV-induced apoptosis. *a*, Immunoblot analysis showing expression of 26K human Bcl-2 protein in pZipBcl-2-transfected AT-3 clones (lanes 4 and 5). Lanes 2 and 3 represent control pZipNeo-transfected AT-3 clone and lane 1 represents wild-type AT-3 cells. *b*, Graph showing cell viability determined by trypan-blue exclusion criteria at serial time points after SV infection (MOI=1) of mock-infected AT3Neo cells (open circles) and AT3Bcl2 cells (open squares) and SV-infected AT3Neo cells (filled circles) and AT3Bcl2 cells (filled squares). Cell viability of each cell type per time point was determined in triplicate; each data point represents the mean viability from three independent experiments $\pm$ s.e.m. *c*, DNA fragmentation patterns of SV-infected AT3Bcl2 cells (lane 2) and SV-infected AT3Neo cells (lane 3) 24 h after infection. Lane 1 shows a 123-bp ladder. *d–g*, Immunofluorescent staining of SV proteins 24 h after infection in mock-infected AT3Neo cells (*d*), SV-infected AT3Neo cells (*e*), mock-infected AT3Bcl2 cells (*f*), and SV-infected AT3Bcl2 cells (*g*), using a primary polyclonal rabbit anti-SV antibody. *h* and *i*, Electron micrographs of SV-infected AT3Neo cells (*h*) and SV-infected AT3Bcl2 cells 24 h after infection. Scale bar, 5 μ m. **METHODS.** A *bcl-2*-expression plasmid was prepared (using standard recombinant DNA methods) by cloning a 910-bp *Eco*RI cDNA from the plasmid pB4 (ref. 25) into the *Bam*HI site of the retroviral expression vector pZipNeo²⁶. The resulting recombinant plasmid pZipBcl-2 and the parental vector pZipNeo were transfected into AT-3 cells by calcium phosphate precipitation and G418 antibiotic-resistant clones were derived as described²⁷. Immunoblot analysis of pZipBcl-2-derived Bcl-2 protein was as described, using antibodies specific for the human Bcl-2 protein²⁸.



with wild-type SV (strain AR339) and examined by videomicroscopy for 72 h. Beginning 18 h after infection, individual cells were seen to undergo a stereotypic pattern of morphological changes in an asynchronous manner (Fig. 1a). The initial event was loss of cell-cell contact, which was followed by nuclear condensation, membrane blebbing, cytoplasmic condensation, and then finally by cell fragmentation and death. By 48 h after infection, nearly all cells appeared to be dead (Fig. 1b). Electron microscopic analysis of BHK-21, AT-3 and N18 cells at 24 h after SV infection (Fig. 1c–e) confirmed the presence of apoptotic ultrastructural morphology. Cells showed early apoptotic features, including intense perinuclear chromatin condensation, loss of microvilli, cytoplasmic vacuolization, and preserved organelle structure (Fig. 1c, d), as well as later apoptotic features such as fragmentation of the nucleus and dying cell to form clusters of membrane-bound apoptotic bodies (Fig. 1e). In addition, gel electrophoresis of BHK-21, AT-3 and N18 cellular DNA revealed fragmentation of chromatin into 180–200 bp oligonucleosomal bands (Fig. 1f). These morphological changes and DNA fragmentation patterns observed in SV-infected BHK-21, AT-3 and N18 cells are diagnostic of apoptosis.

In contrast to the lytic replication cycle established in continuous cell lines (including neuroblastoma cells), SV establishes persistent productive infection in differentiated primary cultured rodent neurons¹. *In vivo*, wild-type SV causes fatal encephalomyelitis in neonatal mice owing to direct cytotoxic effects, but replicates in adult mouse brain without producing significant cytopathology⁹. In immunocompetent mice, productive replication is terminated by an effective humoral immune response, but severe combined immune deficient (*scid*) mice develop long-term persistent productive infection in neurons¹. The distinct difference in the outcome of productive infection (cell death versus persistence) in the immature compared with the mature nervous system, coupled with our observations that SV kills non-neuronal cells by apoptosis, gives an insight into potential cellular factors that may regulate viral lytic potential. During nervous system development, programmed cell death occurs in >50% of the neuronal population, providing a mechanism for matching the size of an overabundant neuronal pool to the amount of target tissue to be innervated^{10,11}. Expression of *bcl-2* has been detected in mouse brain¹² and human adult neocortex¹³, and overexpression of *bcl-2* in rat sympathetic neurons prevents apoptosis induced by deprivation of nerve growth factor³. Therefore we considered that *bcl-2* expression in post-mitotic neurons might play a part in the prevention of SV lytic replication by blocking virus-induced apoptosis.

First, we evaluated whether an association exists between levels of *bcl-2* expression in cultured neurons and susceptibility to lytic viral replication. Fourteen-day-old embryonic rat dorsal root ganglia (DRG) were explanted and cultured in the presence of the anti-mitotic agent fluorodeoxyuridine-uridine (FUDR) (to eliminate all dividing non-neuronal cells) and nerve growth factor (to allow survival and *in vitro* differentiation of neurons). At the time of explantation *bcl-2* messenger RNA is detectable at low levels, but mRNA levels are significantly increased after 6 weeks in culture (Fig. 2a). As the same number of dorsal root ganglia were assayed at each time point, the higher levels of *bcl-2* mRNA observed in 6-week-old cultures reflects increased production by the surviving nerve growth factor-dependent differentiated neurons. Survival time of freshly explanted DRG cultures after SV infection ranges from 1 to 4 days, whereas survival time of differentiated 6-week-old neuron-enriched DRG cultures is more than 14 days (Fig. 2b–d). In addition, SV-infection of 6-week-old neuron-enriched DRG cultures does not result in DNA laddering as observed with other cell types (Fig. 1f, lane 9). These data indicate that there is an association between the level of *bcl-2* expression and resistance to SV-induced cell death in cultured neurons.

Then, to determine directly whether elevated levels of *bcl-2* expression are sufficient to block SV-induced apoptosis, we

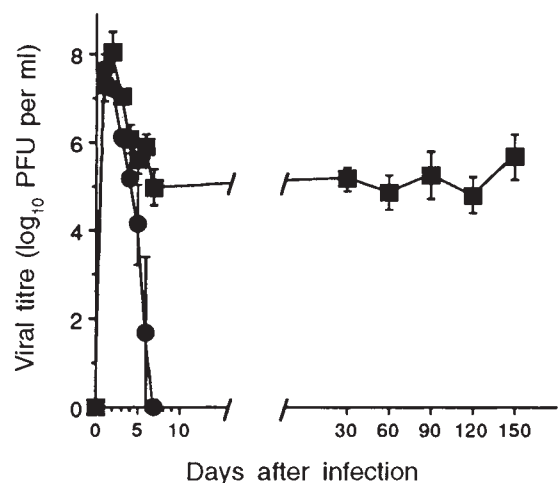


FIG. 4 Growth of SV in AT3Bcl2 cells. Graph showing SV titres in culture supernatants of persistently infected AT3Bcl2 (squares) and acutely infected AT3Neo (circles) cells. Virus titres were measured by plaque assay titration on BHK-21 cells. Each data point represents the geometric mean viral titre from 3–6 experiments $\pm$ s.e.m.

transfected a non-neuronal cell line, AT-3 cells, with the pZipBcl-2 plasmid containing the human *bcl-2* open reading frame and a dominant selectable marker (G418 antibiotic-resistance gene). (AT-3 cells lack any endogenous *bcl-2* mRNA detectable by reverse transcription and polymerase chain reaction; data not shown.) We generated two stable transfectant AT-3 clones expressing human *bcl-2* protein (AT3Bcl2) and two control AT-3 clones transfected with pZipNeo (AT3Neo) (Fig. 3a). Cell viability of AT3Bcl2 and AT3Neo clones was measured at serial time points after SV infection by trypan-blue exclusion (Fig. 3b). The majority of AT3Neo cells were dead by 24 h after infection, and nearly 100% were dead by 48 h after infection. In contrast, the survival of AT3Bcl2 cells after SV infection did not differ significantly from mock-infected AT3Neo or AT3Bcl2 controls. DNA fragmentation into multimers of 180–200 bp was observed in SV-infected AT3Neo cells, but not in SV-infected AT3Bcl2 cells (Fig. 3c). The lack of SV-induced cell death or DNA fragmentation pattern in AT3Bcl2 cells was not due to a block in viral infection or replication. Immunofluorescent staining for SV proteins demonstrated that nearly all AT3Neo and AT3Bcl2 cells were positive at 24 h after infection (Fig. 3d–g), and plaque assay titration of culture supernatants demonstrated that peak viral titres were identical in both cell types (Fig. 4). Electron microscopy of SV-infected AT3Neo cells revealed the same apoptotic morphology seen in SV-infected wild-type AT-3 cells (Fig. 3h). Although SV-infected AT3Bcl2 cells had cytoplasmic vacuoles that are characteristic of alphavirus-infected cells¹⁴, they displayed no nuclear changes or other morphological abnormalities (Fig. 3i). Thus the viability curves, DNA fragmentation patterns, and electron microscopy observations, indicate that AT3Bcl2 cells were protected from SV-induced programmed cell death.

To define the long-term course of SV infection in AT3Bcl2 transfectants, we measured serial viral titres in supernatants from SV-infected AT3Bcl2 cell cultures over a 5-month period (Fig. 4). Early after infection peak viral titres ranged between 10^7 – 10^8 PFU ml⁻¹, and then viral titres decreased by 100–1,000 fold within 6 days after infection. This pattern mimics that observed in *scid* mouse brain¹, and is consistent with previous observations that viral production is less in persistent infection than in acute lytic infection¹⁵. After the initial decrease, viral titres remained constant at 10^4 – 10^6 PFU ml⁻¹ in supernatants from all SV-infected AT3Bcl2 cells followed up to five months after infection. Thus, transfection of AT-3 cells with the *bcl-2*

oncogene transforms an acute lytic infection into a persistent productive infection.

Previous studies have demonstrated that the baculovirus p35 gene¹⁶, Epstein-Barr virus latent genes^{17,18}, adenovirus E1B gene⁶, and herpes simplex virus neurovirulence gene¹⁹ protect cells from apoptosis, suggesting that viral genes may act to promote their own intracellular persistence by enhancing cell survival. In contrast, our data demonstrate that the cellular oncogene *bcl-2* promotes viral persistence by counteracting a cell suicide pathway initiated by alphavirus infection. Our data support an important role for virus-induced apoptosis in lytic

replication, extend previous observations that *bcl-2* can suppress apoptosis^{2,17,20}, and identify a novel host cell mechanism for regulating viral lytic potential. In addition, our findings suggest that differential expression of cellular genes involved in preventing physiological cell death may be an important determinant of tissue-specific viral persistence. The predilection of many viruses to establish persistent infections in cells of the nervous system and the immune system may, in part, be related to the expression of *bcl-2* or of other as-yet unidentified functionally homologous cellular genes that promote longevity. □

Received 5 August; accepted 9 December 1992.

- Levine, B. *et al.* *Science* **254**, 856–860 (1991).
- Hockenbery, D., Nunez, G., Millman, C., Schreiber, R. D. & Korsmeyer, S. J. *Nature* **348**, 334–336 (1990).
- Garcia, I., Martinou, I., Tsujimoto, Y. & Martinou, J. C. *Science* **258**, 302–304 (1992).
- Tera, C., Kornbluth, R. S., Pauza, C. D., Richman, D. D. & Carson, D. A. *J. clin. Invest.* **87**, 1710–1715 (1991).
- Laurent-Crawford, A. G. *et al.* *Virology* **185**, 829–839 (1991).
- Rao, L. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **89**, 7742–7746 (1992).
- Wyllie, A. H., Kerr, J. F. R. & Currie, A. R. *Int. Rev. Cytol.* **68**, 251–306 (1980).
- Arends, M. J. & Wyllie, A. H. *Int. Rev. exp. Path.* **32**, 223–254 (1991).
- Johnson, R. T., McFarland, H. F. & Levy, S. E. *J. infect. Dis.* **125**, 257–262 (1972).
- Martin, D. P. & Johnson, E. M. Jr in *Apoptosis: The Molecular Basis of Cell Death* (eds Tomei, L. D. & Cope, F. O.) 247–261 (Cold Spring Harbor Laboratory Press, New York, 1991).
- Server, A. C. & Mobley, W. C. in *Apoptosis: The Molecular Basis of Cell Death* (eds Tomei, L. D. & Cope, F. O.) 263–278 (Cold Spring Harbor Laboratory Press, New York, 1991).
- Negrini, M., Silini, E., Kozak, C., Tsujimoto, Y. & Croce, C. M. *Cell* **49**, 455–463 (1987).
- Hockenbery, D. M., Zutter, M., Hickey, W., Nahm, M. & Korsmeyer, S. J. *Proc. natn. Acad. Sci. U.S.A.* **88**, 6961–6965 (1991).

- Froshauer, S., Kartenbeck, A. & Helenius, A. *J. Cell Biol.* **107**, 2075–2086 (1988).
- Buchmeier, M. J., Welsh, R. M., Dutko, F. J. & Oldstone, M. B. A. *Adv. Immun.* **30**, 275–331 (1980).
- Clem, R. J., Fechtmeier, M. & Miller, L. K. *Science* **254**, 1388–1390 (1991).
- Henderson, S. *et al.* *Cell* **65**, 1107–1115 (1991).
- Gregory, C. D. *et al.* *Nature* **349**, 612–614 (1991).
- Chou, J. & Roizman, B. *Proc. natn. Acad. Sci. U.S.A.* **89**, 3266–3270 (1992).
- Nunez, G. *et al.* *J. Immun.* **144**, 3602–3610 (1990).
- Martikainen, P., Kyprianou, N., Tucker, R. W. & Isaacs, J. T. *Cancer Res.* **51**, 4693–4700 (1991).
- Sorensen, C. M., Barry, M. A. & Eastman, A. J. *natn. Cancer Inst.* **82**, 749–755 (1990).
- Tsujimoto, Y. & Croce, C. *Proc. natn. Acad. Sci. U.S.A.* **83**, 5214–5218 (1986).
- Levine, B. & Griffin, D. E. *J. Virol.* **66**, 6429–6435 (1992).
- Reed, J. C., Cuddy, M., Slabicki, T., Croce, C. M. & Nowell, P. C. *Nature* **336**, 259–261 (1988).
- Cepko, C., Roberts, B. & Mulligan, R. *Cell* **37**, 1063–1062 (1984).
- Trieger, B. & Isaacs, J. J. *J. Urol.* **140**, 1580–1586 (1988).
- Reed, J. C. *et al.* *Cancer Res.* **51**, 6529–6538 (1991).

ACKNOWLEDGEMENTS. We thank M. Veliuona and S. Dalrymple for technical assistance, and M. Dellanoy for help with electron microscopy. This work was supported by grants from the NIH (B.L., J.T.L., J.C.R., D.E.G., J.M.H.) and the Multiple Sclerosis Society (J.M.H.).

Hepatitis B virus transactivator HBx uses a tumour promoter signalling pathway

Alexander S. Kekulé, Ulrich Lauer, Ludwig Weiss, Birgit Lubert & Peter Hans Hofschneider

Max-Planck-Institut für Biochemie, Am Klopferspitz 18a, D-8033 Martinsried, Germany

THE hepatitis B virus (HBV) transactivator protein HBx is enigmatic in that it stimulates a striking variety of promoters which do not share a common cis-regulatory element^{1–5}. As it does not bind to DNA^{6,7}, it has been speculated that HBx acts indirectly through cellular pathways^{4,6–8}. Under certain conditions HBx can have an oncogenic potential, which may be relevant for HBV-associated liver carcinogenesis^{9–11}, but until now the mechanism for transactivation and cell transformation by HBx was unclear. We report here that HBx uses a complex signal transduction pathway for transactivation. An increase in the endogenous protein kinase C (PKC) activator *sn*-1,2-diacylglycerol and the subsequent activation of PKC give rise to activation of the transcription factor AP-1 (Jun–Fos). As a result, HBx transactivates through binding sites for AP-1 and other PKC-dependent transcription factors (AP-2, NF- κ B), thereby explaining the as-yet incomprehensible variety of HBx-inducible genes. As the PKC signal cascade also mediates cell transformation by tumour-promoting agents, the mechanism presented here might account for the oncogenic potential of HBx.

To investigate whether HBx can use the PKC-dependent transcription factor AP-1 (Jun–Fos) for transactivation, we co-transfected the human liver-derived cell line CCL13 with the HBx expression vector pHbX2s and chloramphenicol acetyltransferase (CAT) reporter plasmids (Fig. 1a). Insertion of one or three AP-1 sites into an otherwise unstimulatable minimal promoter resulted in a 21- or 41-fold transactivation, respectively; a mutated AP-1 sequence used as a control was unresponsive (Fig. 1b).

Besides Jun–Fos (here referred to as AP-1), a whole group of related dimeric transcription factors can also bind to the AP-1 consensus sequence (reviewed in ref. 12). To assess whether transactivation of AP-1 sites by HBx is mediated by the authentic Jun–Fos heterodimer, we co-transfected F9 teratocarcinoma cells, which do not contain detectable amounts of either Jun or Fos¹². HBx alone is unable to transactivate in F9 cells, but CAT induction is fully reconstituted after co-transfection with Jun and Fos expression plasmids (Fig. 1c), showing that Jun–Fos can mediate transactivation by HBx; expression of Jun alone, or of Fos plus a mutated Jun, as a control was insufficient for transactivation.

In response to PKC activation, Jun and Fos dimerize to active, DNA-binding AP-1 (ref. 12). To investigate whether HBx causes activation, that is, increases DNA binding of AP-1, we assayed HBx-expressing CCL13 cells for AP-1 binding activity in an electrophoretic mobility shift assay. Stable transfection with pHbX2s gave rise to a 5–10-fold increase in the retardation signal, as compared with the antisense control pHbX2a (Fig. 1d, lanes 4 and 5), showing that HBx triggers activation of AP-1.

Most mitotic stimuli such as phorbol esters and growth factors activate AP-1 through the PKC pathway, but certain agents, like transforming growth factor and cyclic AMP, seem to activate AP-1 by routes independent of PKC^{12,13}. To elucidate the mechanism of AP-1 activation by HBx, cells were treated with the PKC inhibitor H7. The abrogation of HBx-induced AP-1 retardation (Fig. 1d, lanes 5 and 6) suggests that transactivation by HBx might depend on functional PKC.

To test this possibility, we used H7 in co-transfection experiments. Transactivation of CAT reporter plasmids containing binding sites for the PKC-dependent transcription factors AP-1, AP-2 and NF- κ B (Fig. 1a) was blocked by H7 (Fig. 2a); but dexamethasone induction of the mouse mammary tumour virus (MMTV) long terminal repeat (LTR), included as a PKC-independent control, was not impaired. In another experiment an allosteric inhibitor (sphingosine) and two competitive PKC inhibitors (sangivamycin; GF 109203X), as well as the general protein kinase inhibitor staurosporine, all blocked transactivation as effectively as H7, whereas the protein kinase A-specific inhibitor HA1004 had no effect (Fig. 2b).

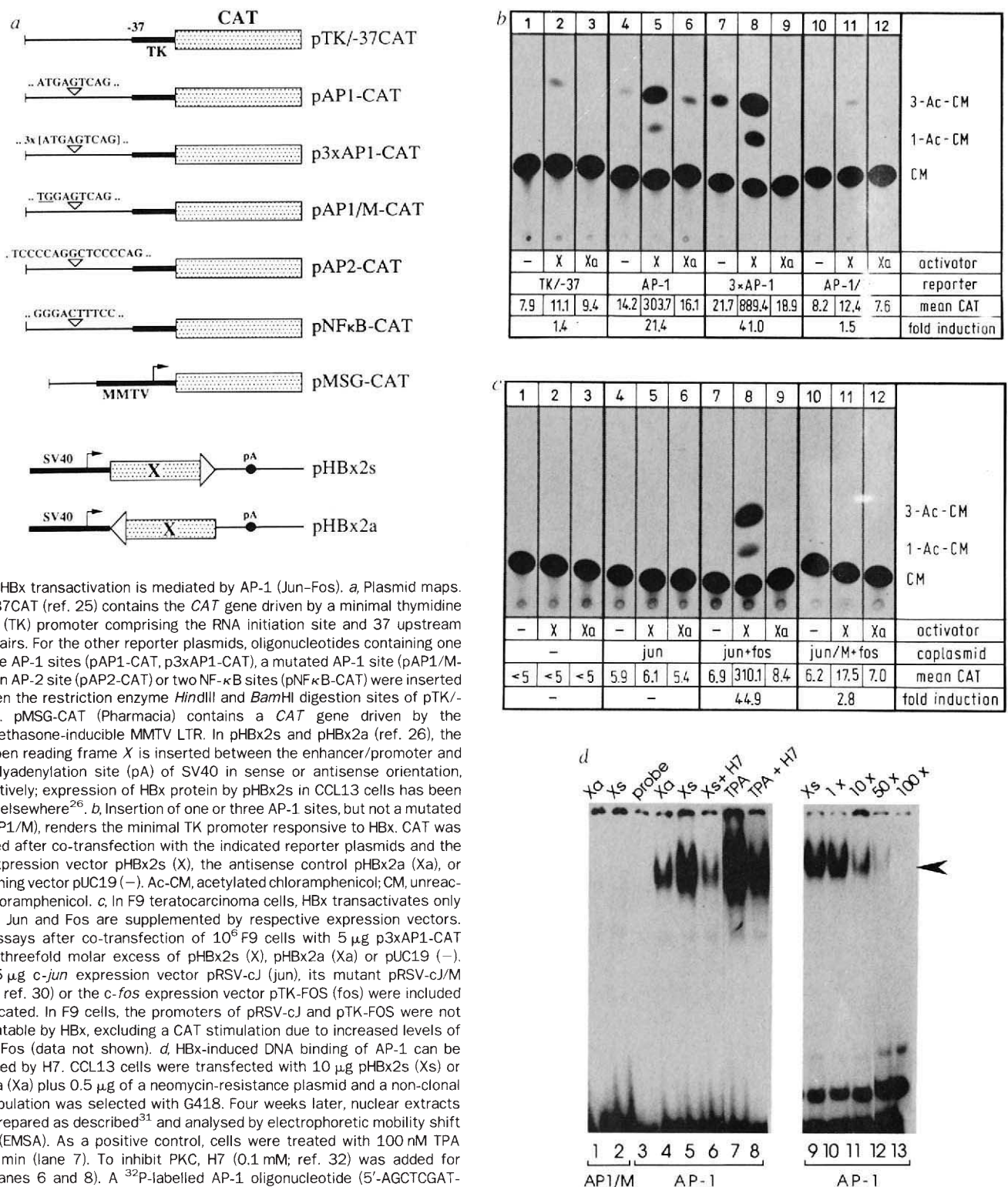


FIG. 1 HBx transactivation is mediated by AP-1 (Jun-Fos). **a**, Plasmid maps. pTK/-37CAT (ref. 25) contains the CAT gene driven by a minimal thymidine kinase (TK) promoter comprising the RNA initiation site and 37 upstream base pairs. For the other reporter plasmids, oligonucleotides containing one or three AP-1 sites (pAP1-CAT, p3xAP1-CAT), a mutated AP-1 site (pAP1/M-CAT), an AP-2 site (pAP2-CAT) or two NF- κ B sites (pNF κ B-CAT) were inserted between the restriction enzyme *Hind*III and *Bam*HI digestion sites of pTK/-37CAT. pMSG-CAT (Pharmacia) contains a CAT gene driven by the dexamethasone-inducible MMTV LTR. In pHbX2s and pHbX2a (ref. 26), the HBV open reading frame X is inserted between the enhancer/promoter and the polyadenylation site (pA) of SV40 in sense or antisense orientation, respectively; expression of HBx protein by pHbX2s in CCL13 cells has been shown elsewhere²⁶. **b**, Insertion of one or three AP-1 sites, but not a mutated site (AP1/M), renders the minimal TK promoter responsive to HBx. CAT was assayed after co-transfection with the indicated reporter plasmids and the HBx expression vector pHbX2s (X), the antisense control pHbX2a (Xa), or the cloning vector pUC19 (-). Ac-CM, acetylated chloramphenicol; CM, unreacted chloramphenicol. **c**, In F9 teratocarcinoma cells, HBx transactivates only if both Jun and Fos are supplemented by respective expression vectors. CAT assays after co-transfection of 10^6 F9 cells with 5 μ g p3xAP1-CAT and a threefold molar excess of pHbX2s (X), pHbX2a (Xa) or pUC19 (-). Also, 5 μ g *c-jun* expression vector pRSV-cJ (jun), its mutant pRSV-cJ/M (jun/M; ref. 30) or the *c-fos* expression vector pTK-FOS (fos) were included as indicated. In F9 cells, the promoters of pRSV-cJ and pTK-FOS were not stimutable by HBx, excluding a CAT stimulation due to increased levels of Jun or Fos (data not shown). **d**, HBx-induced DNA binding of AP-1 can be abolished by H7. CCL13 cells were transfected with 10 μ g pHbX2s (Xs) or pHbX2a (Xa) plus 0.5 μ g of a neomycin-resistance plasmid and a non-clonal cell population was selected with G418. Four weeks later, nuclear extracts were prepared as described³¹ and analysed by electrophoretic mobility shift assay (EMSA). As a positive control, cells were treated with 100 nM TPA for 15 min (lane 7). To inhibit PKC, H7 (0.1 mM; ref. 32) was added for 12 h (lanes 6 and 8). A ³²P-labelled AP-1 oligonucleotide (5'-AGCTCGAT-GAGTCAGCGG-3') was used as probe (lanes 3-13), a mutated AP-1 site (5'-AGCTCGTGGAGTCAGCGG-3') as a negative control (AP1/M; lanes 1, 2). Labelled oligonucleotide (0.5 ng) and 5 μ g protein were incubated in a final volume of 20 μ l (in 10 mM HEPES-NaOH, pH 7.9, 3 mM Tris-HCl, pH 7.9, 60 mM KCl, 30 mM NaCl, 0.2 mM EDTA, 5 mM MgCl₂, 1 mM dithiothreitol, 10% (v/v) glycerol, 2 μ g poly(dI-dC), 10 μ g bovine serum albumin) and separated on a low-ionic-strength 4% polyacrylamide gel (arrow indicates position of the AP-1 retardation signal). Specificity for AP-1 was confirmed by extinction of the retardation signal after addition of 1- to 100-fold excess (lanes 10-13) of non-labelled AP-1 oligonucleotide to nuclear extracts of HBx-transfected cells (lane 9: without competitor).

METHODS. 6×10^5 CCL13 cells (American Type Culture Collection) were seeded into 6-cm dishes and co-transfected 24 h later with 5 μ g reporter and a threefold molar excess of activator plasmid by the calcium phosphate

co-precipitation method, applying a glycerol shock after 3 h (ref. 27). Transfection efficiencies were standardized with a luciferase reporter plasmid containing the luciferase gene²⁸ under the control of the non-stimulatable minimal promoter of pTK/-37CAT. Inhibitors were added to the culture medium immediately after the glycerol shock. Cell extracts were prepared 48 ± 5 h later and protein concentrations determined by a commercial kit (Bio-Rad). CAT was assayed as described²⁹ with 50-100 μ g of protein and acetylated and unreacted ¹⁴C-chloramphenicol quantified with an isotope scanner (Berthold, Germany). Mean CAT activities were calculated from three independent transfections (in pmol Ac-CM per min per mg protein, standard deviation 5-15%) and therefore do not fully agree with the autoradiograph in all lanes. Fold induction: ratio of the stimulated (X) to the control values (-).

To demonstrate PKC dependence unequivocally, we took advantage of the fact that eukaryotic cells can be specifically depleted of PKC by long-term treatment with 12-*O*-tetradecanoyl-phorbol-13-acetate (TPA)¹⁴. After co-transfection of PKC-depleted CCL13 cells, no more transactivation by HBx could be observed (Fig. 2c); in contrast, dexamethasone induction of the mouse MMTV LTR and ultraviolet activation of the *c-fos* promoter, which are both independent of PKC¹⁵, were not affected.

Taken together, these results demonstrate that HBx requires functional PKC for transactivation. To test whether HBx actually causes activation of PKC, namely translocation of the inactive cytosolic enzyme to the cell membrane¹⁴, we fractionated transiently transfected CCL13 cells. Whereas resting or control-transfected cells contained virtually no PKC activity in the membrane fraction, expression of HBx caused a significant translocation of PKC to the membrane (Fig. 3a). Activation was transient, with a maximum 5 h post-transfection and full reversion after 24 h (data not shown), possibly explaining why

in another study¹⁶ no PKC translocation was detected 36 h after infection with an HBx-expressing recombinant adenovirus.

Most physiological stimuli (such as membrane receptors, Ras, Src and G proteins) elicit the PKC pathway indirectly by increasing the endogenous PKC activator 1,2-diacylglycerol (DAG)¹⁴. To determine whether HBx also influences DAG, growth-arrested NIH3T3 cells, containing a lower basal DAG level than CCL13, were transfected with pHbX2s. Subsequent quantitation of DAG revealed a 1.9-fold elevation in response to HBx expression (fig. 3b), strongly suggesting that HBx-dependent PKC activation is mediated by DAG.

In conclusion, HBx can use a complex signalling pathway for transactivation, comprising DAG, PKC, AP-1 and probably other PKC-dependent transcription factors such as AP-2 and NF- κ B (Fig. 4). This mechanism of gene regulation, which is unprecedented among viral transactivators, resembles the effects of G proteins, and Src and Ras proteins, which likewise elicit PKC through DAG¹⁷⁻²⁰. However, the mechanism by which HBx causes an increase in DAG is still unclear. Both a cyto-

FIG. 2 Transactivation by HBx is dependent on protein kinase C. *a*, Transactivation of p3xAP1-CAT, pAP2-CAT and pNF κ B-CAT by pHbX2s is blocked by H7 (0.1 mM; lanes 1-9). Induction of the MMTV LTR in pMSG-CAT by 10⁻⁷ M dexamethasone (dx) is not impaired (lanes 10-13). For methods and abbreviations, see Fig. 1 legend. *b*, Transactivation is blocked by PKC inhibitors H7 (0.1 mM), sangivamycin³³ (0.1 mM), sphingosine³⁴ (0.02 mM), GF 109203X (GF, 2 μ M; ref. 35) and staurosporine (stauro, 0.1 μ M; ref. 36), whereas the PKA-specific inhibitor HA 1004 (0.1 mM; ref. 37) has no significant effect. CAT assays after co-transfection of CCL13 cells with p3xAP1-CAT and pHbX2s (x) or pUC19 (-), respectively; lanes 1-3, without inhibitor. To monitor PKC inhibition at the applied inhibitor concentrations, p3xAP1-CAT was induced with 50 nM TPA (T). Inhibitor concentrations match >90% PKC inhibition of dose-response relationships and are below toxic doses determined for CCL13 cells. *c*, HBx does not transactivate in CCL13 cells depleted of PKC by 36 h treatment with 300 nM TPA (depletion confirmed by quantitative PKC assay). CAT was assayed after co-transfection of p3xAP1-CAT with pHbX2s (x), pUC19 (-) or pUC19 plus treatment with 50 nM TPA (T). As PKC-independent controls, cells were transfected with 2 μ g pMSG-CAT (MMTV; lanes 7-8) or 15 μ g *pcfos*(-711)CAT³⁸ (*fos*; lanes 9-10) and stimulated with 10⁻⁷ M dexamethasone (dx) or by ultraviolet irradiation¹⁵, respectively (254 nm, ~30 J m⁻²).

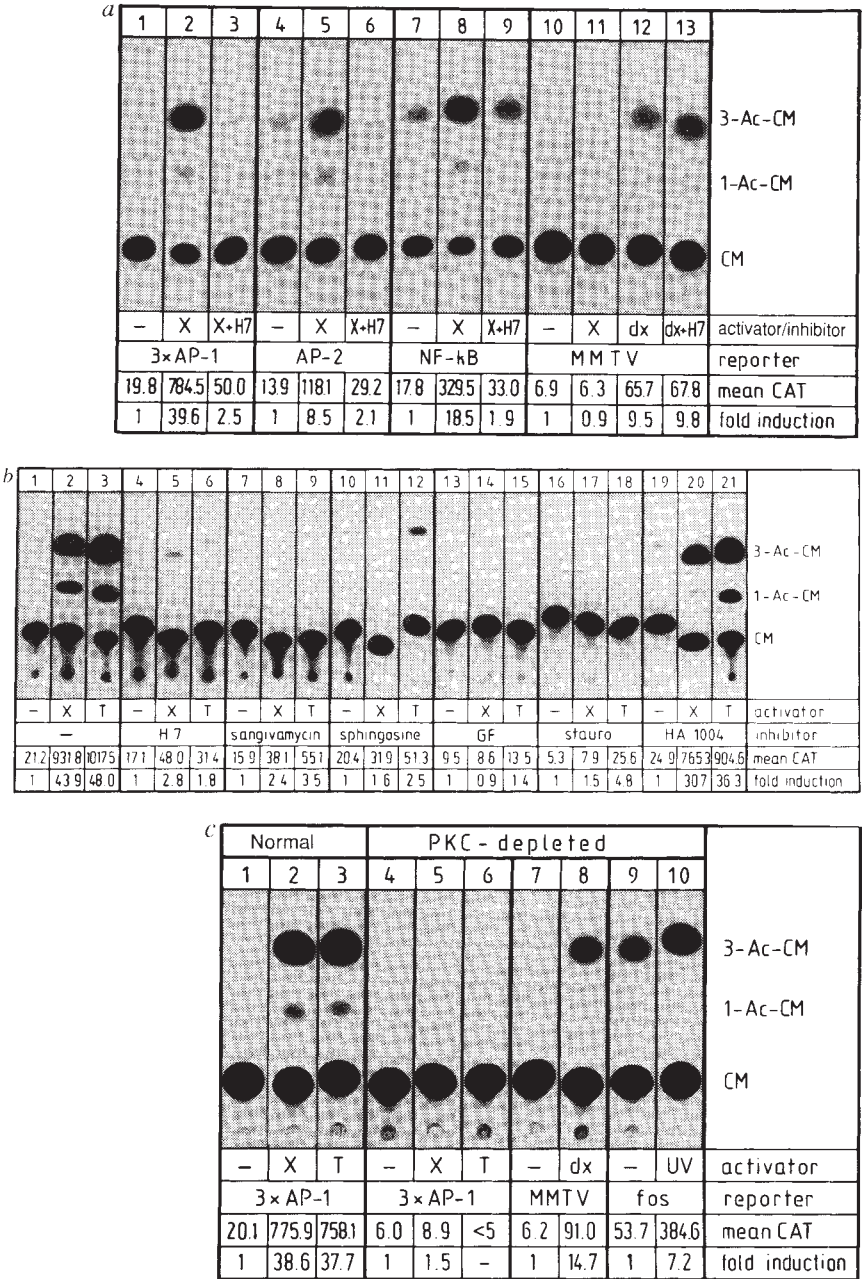
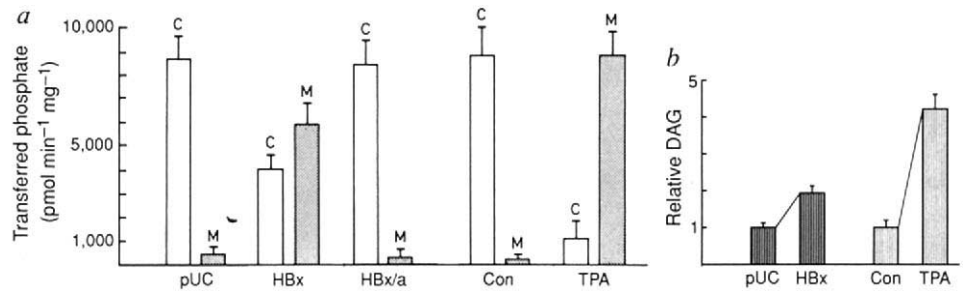


FIG. 3 Expression of HBx causes membrane translocation of PKC and elevation of the endogenous PKC activator DAG. *a*, Translocation of PKC from the cytosol (open bars, C) to the membrane (shaded bars, M). CCL13 cells were growth-arrested by 36 h low-serum treatment (0.5%) and transfected with pUC19 (pUC), pHbX2s (HBx) or pHbX2a (HBx/a). Membrane and cytosol fractions were prepared 5 h post-transfection. PKC was partially purified as described³⁹ and determined both by phosphorylation of a specific substrate (Amersham) and the histone H1 mixed micelle assay³⁹. Con, untransfected growth-arrested cells; TPA, stimulation with 800 nM TPA for 10 min. Shown are the mean values of three experiments, in pmol transferred phosphate per min per mg protein. *b*, Transfection with pHbX2s (HBx) causes a 1.9-fold DAG elevation as compared to pUC19 (pUC). As a control, untransfected cells (Con) were stimulated with TPA (1 μ M)¹⁷. 1.5×10^6 NIH3T3 cells were



growth-arrested by 24 h low-serum treatment (0.5%) and lipofected according to the manufacturer's recommendations (DOTAP; Boehringer Mannheim). Cells were collected 4 h later and phospholipid concentrations standardized⁴⁰. DAG was determined by quantitative conversion with *E. coli* DAG kinase⁴¹. Shown are mean DAG values of three transfections relative to pUC-transfected and untransfected controls, respectively.

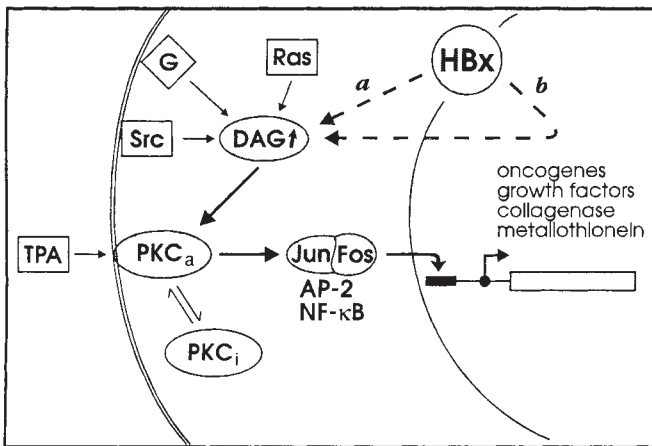


FIG. 4 HBx elicits a complex pathway comprising elevation of DAG and translocation of cytosolic PKC (inactive, PKC_i) to the membrane (active, PKC_a), followed by target gene activation by AP-1 (Jun-Fos) and probably also other PKC-dependent transcription factors (AP-2, NF- κ B). DAG elevation by HBx, which is present in the cytoplasm and the nucleus, could be due to a cytoplasmic (*a*) or a nuclear effect (*b*).

plasmic effect, for example as a result of its presumed protein kinase activity⁸, and a more indirect interaction with nuclear transcription factors (like CREB/ATF-2; ref. 21) would be consistent with our results (*a* and *b* in Fig. 4).

As replication of animal viruses is often activated by PKC-stimulating agents such as TPA and growth factors¹⁵, excitation of the PKC pathway by HBx may be necessary for the production of large amounts of virus particles. This view is supported by the finding that HBV mutants defective for HBx expression can still replicate, albeit with reduced efficiency^{22,23}.

As HBx transforms immortalized mouse hepatocytes^{9,11} and causes tumours in transgenic mice¹⁰, it may be involved in HBV-associated hepatocarcinogenesis^{1-3,9-11}. But its inability to transform NIH3T3 cells (ref. 10, and our unpublished observations) and the long latency preceding carcinoma development have suggested that its oncogenic activity is different from other known oncoproteins¹¹. Interestingly, the PKC signalling pathway mediates tumour promotion by growth factors, mitogens and phorbol esters¹⁵; accordingly, overproduction of PKC causes disordered growth in mammalian cells²⁴. Therefore the finding that a viral gene product can also activate this tumour promoter pathway might account for the atypical oncogenicity of HBx. □

Received 28 September; accepted 3 December 1992.

- Zahn, P., Hofschneider, P. H. & Koshy, R. *Oncogene* **3**, 169-177 (1988).
- Tsu, J. & Schloemer, R. H. *J. Virol.* **61**, 3448-3453 (1987).
- Spandau, D. F. & Lee, C. *J. Virol.* **62**, 427-434 (1988).
- Seto, E., Mitchell, P. J. & Yen, T. B. *Nature* **344**, 72-74 (1990).
- Rossner, M. T. *J. med. Virol.* **36**, 101-117 (1992).
- Faktor, O. & Shaul, Y. *Oncogene* **5**, 867-872 (1990).
- Siddiqui, A., Jameel, S. & Mapoles, J. *Proc. natn. Acad. Sci. U.S.A.* **84**, 2513-2517 (1987).

- Wu, J. Y., Zhuo, Z.-Y., Judd, A., Cartwright, C. A. & Robinson, W. S. *Cell* **63**, 687-695 (1990).
- Höhne, M. et al. *EMBO J.* **9**, 1137-1145 (1990).
- Kim, C.-M., Koike, K., Saito, I., Miyamura, T. & Jay, G. *Nature* **351**, 317-320 (1991).
- Seifer, M., Höhne, M., Schaefer, S. & Gerlich, W. H. *J. Hepatology* **13**, S61-S65 (1991).
- Angel, P. & Karin, M. *Biochim. biophys. Acta* **1072**, 129-157 (1991).
- Müller, U., Roberts, M. P., Engel, D. A., Doerfler, W. & Shenk, T. *Genes Dev.* **3**, 1991-2002 (1989).
- Nishizuka, Y. *Nature* **334**, 661-665 (1988).
- Büscher, M., Rahmsdorf, H. J., Litfin, M., Karin, M. & Herrlich, P. *Oncogene* **3**, 301-311 (1988).
- Lucito, R. & Schneider, R. J. *J. Virol.* **66**, 983-991 (1992).
- Exton, J. H. *J. biol. Chem.* **265**, 1-4 (1990).
- Lacal, J. C., Joscot, J. & Aaronson, S. A. *Nature* **220**, 269-272 (1987).
- Song, J., Pfeffer, L. M. & Foster, D. A. *Molec. cell. Biol.* **11**, 4903-4908 (1991).
- Wasyluk, C., Imier, J. L., Perez-Mutul, J. & Wasyluk, B. *Cell* **48**, 525-534 (1987).
- Maguire, H. F., Hoeffer, J. P. & Siddiqui, A. *Science* **252**, 842-844 (1991).
- Blum, H. E. et al. *J. Virol.* **66**, 1223-1227 (1992).
- Yaginuma, K., Shirakata, Y., Kobayashi, M. & Koike, K. *Proc. natn. Acad. Sci. U.S.A.* **84**, 2678-2682 (1987).
- Housey, G. M. et al. *Cell* **52**, 343-354 (1988).
- Baniahmad, A., Steiner, C., Köhne, A. C. & Renkawitz, R. *Cell* **61**, 505-514 (1990).
- Wollersheim, M. & Hofschneider, P. H. in *Viral Hepatitis and Liver Disease* (ed. Zuckerman, A. J.) 334-340 (Liss, New York, 1988).
- Graham, F. L. & van der Eb, A. J. *Virology* **52**, 456-467 (1973).
- De Wet, J. R., Wood, K. V., DeLuca, M., Helinski, D. R. & Subramani, S. *Molec. cell. Biol.* **7**, 725-737 (1987).
- Gorman, C. M., Moffat, L. F. & Howard, B. H. *Molec. cell. Biol.* **2**, 1044-1051 (1982).
- Angel, P., Hattori, K., Smeal, T. & Karin, M. *Cell* **55**, 875-885 (1988).
- Dignam, J. P., LeBovitz, R. M. & Roeder, R. G. *Nucleic Acids Res.* **11**, 1475-1489 (1983).
- Hidaka, H., Inagaki, M., Kawamoto, S. & Sasaki, Y. *Biochemistry* **23**, 5036-5041 (1984).
- Loomis, C. R. & Bell, R. M. *J. biol. Chem.* **263**, 1682-1692 (1988).
- Hannun, Y. A., Loomis, C. R., Merrill, A. H. & Bell, R. M. *J. biol. Chem.* **261**, 12604-12609 (1986).
- Toullec, D. et al. *J. biol. Chem.* **266**, 15771-15781 (1991).
- Tamaoki, T. et al. *Biochem. biophys. Res. Commun.* **135**, 397-402 (1986).
- Asano, T. & Hidaka, H. *J. Pharmacol. exp. Ther.* **231**, 141-145 (1984).
- König, H. et al. *EMBO J.* **8**, 2559-2566 (1989).
- Fabbro, D., Jungmann, R. A. & Eppenberger, U. *Arch. Biochem. Biophys.* **239**, 102-111 (1985).
- Van Veldhoven, P. P. & Mannaerts, G. P. *Analyt. Biochem.* **161**, 45-48 (1987).
- Preiss, J. et al. *J. biol. Chem.* **261**, 8597-8600 (1986).

ACKNOWLEDGEMENTS. We thank M. Karin and P. Angel for *c-jun* expression vectors; A. Baniahmad and C. Steiner for plasmids pTK-FOS and pTK-37CAT; H. J. Rahmsdorf for plasmid *pcfos*/CAT; Gödecke GmbH for synthesis of PKC inhibitor GF 109203X; H. Grunike, P. Herrlich and R. Koshy for discussion and critical reading of the manuscript; F. Überall and G. Sauter for help with the PKC and DAG assays; and A. Schultz, Z. Squire and C. Schommer for technical assistance. This work was supported by the Deutsche Stiftung für Krebsforschung, Bonn.

Local domains of supercoiling activate a eukaryotic promoter *in vivo*

Marietta Dunaway* & Elaine A. Ostrander†

* Division of Biochemistry and Molecular Biology, Department of Molecular and Cell Biology, University of California, Berkeley, California 94720, USA

† Department of Cellular and Molecular Biology, Human Genome Center, Lawrence Berkeley Laboratory, Berkeley, California 94720, USA

EXPERIMENTS correlating template topology with transcriptional activity suggest that DNA topology plays a role in eukaryotic gene expression¹⁻⁴. Linear templates transfected into cultured cells produce far fewer transcripts than do circular transcription templates³, and no transcripts can be detected from linear templates injected into *Xenopus* oocytes^{1,2}. Further, when transcriptionally active circular templates in *Xenopus* oocytes are linearized by injection of a restriction enzyme, transcription dramatically decreases. Here we show that transcription by phage T7 RNA polymerase from a divergent promoter can partially replace the requirement for circular *Xenopus* ribosomal RNA transcription templates in *Xenopus* oocytes. Supercoiled domains can apparently be generated on short pieces of DNA having no known sequences that result in association with the nuclear architecture, suggesting that localized, transient domains of supercoiling fulfil the minimum topological needs for *Xenopus* rRNA transcription *in vivo*.

If the transcription machinery cannot rotate freely around the DNA during transcriptional elongation, two distinct domains of localized supercoiling result⁵ (Fig. 1A). Positive supercoils form ahead of the advancing RNA polymerase while negative supercoils accumulate in the DNA behind it. When the supercoils are prevented from diffusing by tethering of the DNA template, distinct supercoiled domains will be created around actively transcribed genes (Fig. 1A). Supercoiling resulting from transcriptional elongation has been demonstrated both *in vitro* and *in vivo*⁵⁻⁹. The physiological significance of the relationship between supercoiling and transcription, however, remains unresolved.

Templates were constructed containing both a phage T7 promoter and the *Xenopus* rRNA promoter (Fig. 1B). The T7 promoter was inserted in a divergent orientation upstream (template b), or in a tandem orientation downstream (template c) with respect to the *Xenopus* rRNA promoter. In both templates transcription by T7 RNA polymerase will potentially create negative supercoils at the rRNA promoter (Fig. 1A). Linearized templates and T7 RNA polymerase were coinjected into *Xenopus* oocytes, and the rRNA transcripts from the injected templates were detected using an S1 nuclease assay^{10,11}.

As expected, none of the linear DNAs was transcribed in injected oocytes (Fig. 2A, lanes 3-6). But when the linear templates were coinjected with T7 RNA polymerase, rRNA transcription was activated in template b (Fig. 2A, lane 9), where the T7 promoter is placed immediately upstream and in divergent orientation with respect to the rRNA promoter. Coinjection of the polymerase with transcription templates c and d did not result in detectable rRNA transcription under these conditions. Although T7 transcription decreases template stability some-

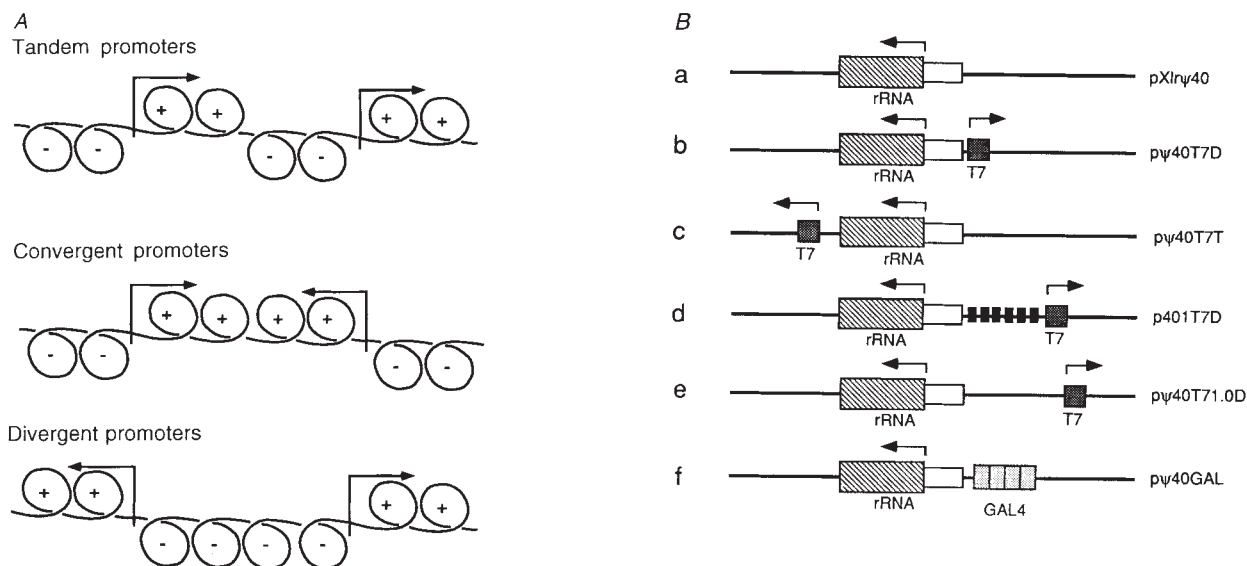


FIG. 1 Experimental strategy. **A**, Predicted supercoiled domains between two promoters in three possible relative orientations. **B**, Templates (a-f) used in these studies. **a**, All plasmids for the *Xenopus* rRNA genes are based on the ψ 40 minigene¹⁵ cloned into the vector pUC118. The minigene contains the full promoter region ($\sim$ 245), approximately 110 bases of the external transcribed spacer, and the last 200 bases of the 28S region, including 160 bases after the processed 3' end of the transcript. The white rectangle symbolizes the promoter. The minigenes pXlr ψ 40 and pXlr ψ 52 contain 40 or 52-bp insertions, respectively; therefore transcripts from these templates can be distinguished from each other and from the endogenous transcripts by S1 analysis. **b**, A *PvuII*-*SphI* fragment containing the promoter for T7 RNA polymerase from pGEM3 was inserted directly into the ψ 40 minigene in a divergent orientation to create p ψ 40T7D. **c**, The T7 promoter from a Bluescript plasmid was used to make the tandem orientation

of the T7 promoter in p ψ 40T7T. **d**, Plasmid p401T7D was constructed as described for template b, beginning with a minigene containing the enhancer (pXlr401; ref. 15). Black boxes symbolize the tandem array of xUBF binding sites that comprise the enhancer^{16,18}. **e**, Plasmid p ψ 40T7 1.0D was prepared in the same way as ψ 40T7D, using a *DraI* and *SphI* T7 promoter fragment from pGEM3. The T7 and rRNA initiation sites are separated by $\sim$ 270 bases in template b, 680 bp in template c, 970 bp in template d, and 1,100 bp in template e. **f**, p ψ 40GAL was made by inserting four GAL4 17-bp consensus binding sites at the same position that the T7 promoter occupies in p ψ 40T7D. Templates are not drawn to scale. Not all relative orientations of the T7 and rRNA promoters could be tested because T7 RNAP transcription disrupted the rRNA transcription complex (data not shown), or because T7 RNAP transcription interfered with the S1 assay for *Xenopus* rRNA transcripts.

what, all of the injected templates survive the time course of the experiment and no differential template stability was observed (Fig. 2B: compare lanes 2 and 3 to lanes 7 and 8). When a circular reference template is coinjected with linear template b and T7 RNA polymerase, rRNA transcription from the linear template is neither diminished nor increased (Fig. 2C). The circular template therefore does not titrate either inhibitory or stimulatory factors, suggesting that the two templates behave in a similar manner.

Coinjection of T7 RNA polymerase with linear template d did not activate *Xenopus* rRNA transcription, even though template d contains T7 and rRNA promoters in divergent orientation and includes the rRNA enhancer (Fig. 2A, lane 11). The rRNA promoter was not activated in template e either (Fig. 4, lanes 2–4), suggesting that rRNA transcriptional activation by T7 RNA polymerase coinjection is sensitive to the distance between the two promoters.

To determine whether rRNA transcription on linear templates is dependent on T7 RNA polymerase transcription, we first tested the dependence of transcription from both promoters on polymerase concentration. When increasing concentrations of

T7 RNA polymerase were coinjected with linear template b, both rRNA transcription and transcripts produced by the T7 RNA polymerase increased concomitantly (Fig. 3A and B). To exclude the possibility that proximity of the polymerase protein surface or protein mass to the rRNA promoter facilitates rRNA transcription, we coinjected a chimaeric protein composed of the DNA-binding domain of the transcriptional activator GAL4 fused to T7 RNA polymerase (ref. 8) with templates b and f. This chimaeric protein binds to the 17-base-pair (bp) consensus GAL4 binding site and can initiate transcription from a T7 promoter. The GAL4/T7 RNA polymerase fusion protein activated rRNA transcription from template b (Fig. 3c, lane b), but not linear template f (lane f). As protein binding but not transcriptional initiation is permitted by the GAL4 binding sites in template f, we conclude that T7 transcription and not the proximity of T7 RNA polymerase is required to activate rRNA transcription on template b.

If rRNA transcription is activated by supercoiling resulting from T7 RNA polymerase transcription, then inhibitors of endogenous topoisomerases should increase supercoiling and therefore increase transcription of the rRNA promoter. To test

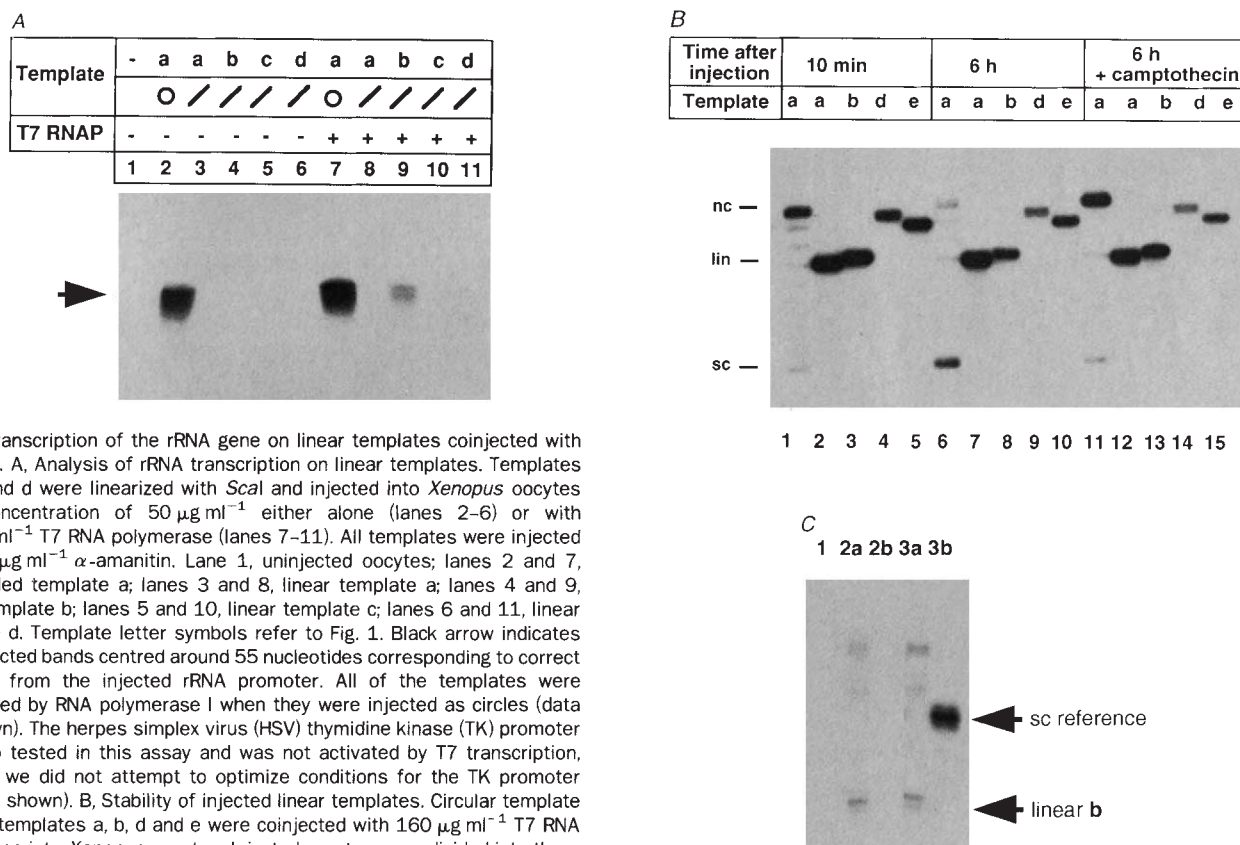


FIG. 2 Transcription of the rRNA gene on linear templates coinjected with T7 RNAP. A, Analysis of rRNA transcription on linear templates. Templates a, b, c and d were linearized with *ScaI* and injected into *Xenopus* oocytes at a concentration of $50 \mu\text{g ml}^{-1}$ either alone (lanes 2–6) or with $160 \mu\text{g ml}^{-1}$ T7 RNA polymerase (lanes 7–11). All templates were injected with $50 \mu\text{g ml}^{-1}$ α -amanitin. Lane 1, uninjected oocytes; lanes 2 and 7, supercoiled template a; lanes 3 and 8, linear template a; lanes 4 and 9, linear template b; lanes 5 and 10, linear template c; lanes 6 and 11, linear template d. Template letter symbols refer to Fig. 1. Black arrow indicates S1-protected bands centred around 55 nucleotides corresponding to correct initiation from the injected rRNA promoter. All of the templates were transcribed by RNA polymerase I when they were injected as circles (data not shown). The herpes simplex virus (HSV) thymidine kinase (TK) promoter was also tested in this assay and was not activated by T7 transcription, although we did not attempt to optimize conditions for the TK promoter (data not shown). B, Stability of injected linear templates. Circular template a, linear templates a, b, d and e were coinjected with $160 \mu\text{g ml}^{-1}$ T7 RNA polymerase into *Xenopus* oocytes. Injected oocytes were divided into three groups. Lanes 1–5 show the injected DNA from oocytes isolated immediately after injection was completed, and lanes 6–10, the DNA from oocytes collected after 6 h. The third group of oocytes (lanes 11–15), was incubated in buffer containing $100 \mu\text{g ml}^{-1}$ camptothecin for 6 h and the oocytes were then collected. Total nucleic acid was purified from all oocytes, and one oocyte-equivalent from each sample was electrophoresed through an agarose gel, the DNA transferred to GeneScreen and hybridized with a probe for the rRNA minigene. Bands therefore correspond to the molar ratios of the DNAs. Lanes 1, 6, 11: circular template a; lanes 2, 7, 12: linear template a; lanes 3, 8, 13: linear template b; lanes 4, 9, 14: linear template d; lanes 5, 10, 15: linear template e. There is no evidence of either concatemerization or recircularization, consistent with previous results^{11,17}. nc, Nicked circle; lin, linear; sc, supercoiled. C, Competition between circular and linear rRNA transcription templates. Linear template b was injected into oocytes alone, (sample 1), with T7 RNAP (sample 2; lanes 2a and 2b), and with T7 RNAP plus the circular rRNA template $\psi 52$ (sample 3; lanes 3a and 3b). Sample 1 was assayed for the $\psi 40$ transcript, and samples 2 and 3 were assayed in parallel with specific probes for the $\psi 40$ transcript (lanes 2a and 3a) and

for circular $\psi 52$ (lanes 2b and 3b). Correct transcription initiation for each template is indicated by labelled black arrows. We observed aberrant upstream starts (open arrows) in some experiments that vary with oocyte batch.

METHODS. Each oocyte was injected with 20 nl of a mix containing the specific components. At least 50 oocytes were injected for each sample, and experiments were analysed only when oocyte survival was $>50\%$. Transcripts were allowed to accumulate for 6 h, oocytes were collected and lysed in 100 mM NaCl, 50 mM Tris-Cl, pH 8.0, 10 mM EDTA, 1% SDS, and total nucleic acid was purified. Transcripts from injected templates were detected using a probe specific for the template in an S1 nuclease assay¹¹. The kinase-labelled 5' end of the single-stranded probe specific for $\psi 40$ or $\psi 52$ corresponds to the end of the insertions in the minigene and extends to the *BglI* site at -65 . Thus each template-specific probe hybridizes with the endogenous 40S precursor and with both minigene transcripts, but produces a labelled protected fragment only when the 5' end of the S1 probe exactly matches the transcript.

this prediction we used camptothecin, an inhibitor of topoisomerase I (ref. 12). The three linearized constructs b, d and e were coinjected with T7 RNA polymerase. Half of the oocytes from each injection were incubated with camptothecin and assayed as described above. In the presence of camptothecin, transcription from the linear rRNA template b increased fivefold (Fig. 4, lanes 2 and 5). Moreover, transcripts from templates d and e were not easily detected (lanes 6 and 7). The stability of the linear templates is not affected by camptothecin (Fig. 2B; compare lanes 6–10 with lanes 11–15), and the high proportion of circular plasmid running at the nicked position indicates that most templates have an inhibited topoisomerase I molecule associated with them. Activation of rRNA transcription on linear templates is therefore apparently related to changes in template supercoiling, suggesting that supercoiling generated by T7 transcription activates the rRNA promoter.

These results suggest a mechanism for both the source and the role of supercoiling in rRNA transcription. In our experiments transcription-driven supercoils were produced from a nearby divergent promoter, an arrangement that does not occur in rRNA genes. We suggest that the source of transcription-

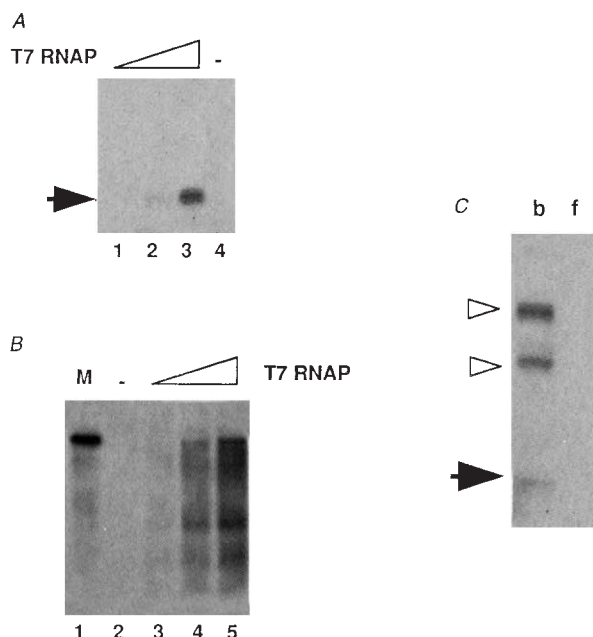


FIG. 3 rRNA transcription on linear templates is dependent on T7 RNA polymerase activity. A, rRNA transcription increases proportionally with T7 RNA polymerase concentration (T7 RNAP). Linear template b was injected into *Xenopus* oocytes with $50 \mu\text{g ml}^{-1}$ α -amanitin and: lane 1, $40 \mu\text{g ml}^{-1}$ T7RNAP; lane 2, $80 \mu\text{g ml}^{-1}$ T7 RNAP; lane 3, $160 \mu\text{g ml}^{-1}$ T7 RNAP; lane 4, no T7 RNAP. B, T7 transcription increases proportionally with T7 RNAP concentration. A portion of the samples analysed in A (lanes 2–5) and the T7 transcript from template b prepared *in vitro* were electrophoresed through a formaldehyde gel. After electrophoresis, the RNA was transferred to GeneScreen, crosslinked by ultraviolet irradiation and hybridized with a plasmid fragment corresponding to the T7 transcript. Lane 1, 0.5 ng *in vitro*-synthesized T7 RNA from template b; lane 2, RNA from oocytes injected with template b; lane 3, template b injected with $40 \mu\text{g ml}^{-1}$ T7 RNAP; lane 4, b with $80 \mu\text{g ml}^{-1}$ T7 RNAP; lane 5, b with $160 \mu\text{g ml}^{-1}$ T7 RNAP. C, *Xenopus* rRNA transcription depends on T7 RNA transcription. Linear templates b and f ($50 \mu\text{g ml}^{-1}$) were injected into oocytes with the GAL4/T7RNAP fusion protein. Lane b, linear template b; lane f, linear template f. Black arrow indicates correct initiation from the rRNA promoter. All templates were also coinjected with the GAL4/T7 RNAP fusion protein. We find that the fusion protein and T7 RNAP are interchangeable in these experiments. But because the GAL4/T7 RNAP chimera also creates a loop, we cannot distinguish whether simple looping or transcription-driven supercoils activate the rRNA promoter. For this reason, most of the experiments reported here use T7 RNAP.

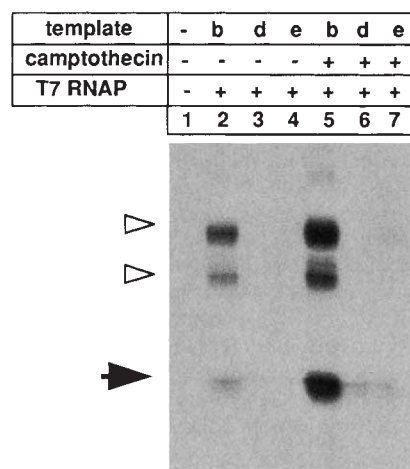


FIG. 4 Camptothecin stimulates rRNA transcription on linear templates. Linear templates were coinjected into *Xenopus* oocytes with $160 \mu\text{g ml}^{-1}$ T7 RNAP and $50 \mu\text{g ml}^{-1}$ α -amanitin. Oocytes from each injection were divided into two sets: one set was incubated for 6 h in buffer supplemented with $100 \mu\text{g ml}^{-1}$ camptothecin (lanes 5–7); the second set was incubated in buffer without camptothecin (lanes 1–4). Oocytes were collected and processed as above. Black arrow indicates the correctly initiated rRNA transcript; arrowheads indicate aberrant starts. Lane 1, uninjected oocytes; lanes 2 and 5, template b; lanes 3 and 6, template d; lanes 4 and 7, template e.

driven supercoiling in the rRNA genes is RNA polymerase I leaving the promoter and producing negative supercoils in its wake. This model predicts a mechanistic distinction between the first transcription initiation event, which would occur without benefit of transcription-driven supercoiling, and subsequent initiations, which could be facilitated by transcription-driven supercoiling. Supercoils could be created for the first initiation event from disassembly of nucleosomes, or the supercoiling deficit in the first initiation event may be compensated by the participation of enhancers and upstream promoter elements in the preinitiation complex.

One condition of the twin-supercoiled-domain model is the requirement for a tether to capture the supercoils generated by the transcribing polymerase⁵. In the chromosome, tethers could be provided by DNA elements known as scaffold-associated regions^{13,14}. Our results suggest that local areas of supercoils can be generated along DNA without attachment to the nuclear architecture, and that these small domains of supercoiling can affect transcription. □

Received 1 November; accepted 30 December 1992.

- Harland, R. M., Weintraub, H. & McKnight, S. L. *Nature* **301**, 38–43 (1983).
- Pruitt, S. C. & Reeder, R. H. *J. molec. Biol.* **174**, 121–139 (1984).
- Weintraub, H., Cheng, P. F. & Conrad, K. *Cell* **46**, 115–122 (1986).
- Schultz, M. C., Brill, S. J., Ju, Q., Sternglanz, R. & Reeder, R. H. *Genes Dev.* **6**, 1332–1341 (1992).
- Liu, L. F. & Wang, J. C. *Proc. natn. Acad. Sci. U.S.A.* **84**, 7024–7027 (1987).
- Glaever, G. N. & Wang, J. C. *Cell* **55**, 849–856 (1988).
- Tsao, Y.-P., Wu, H.-Y. & Liu, L. F. *Cell* **56**, 111–118 (1989).
- Ostrander, E. A., Benedetti, P. & Wang, J. C. *Science* **249**, 1261–1265 (1990).
- Ljungman, M. & Hanawalt, P. C. *Proc. natn. Acad. Sci. U.S.A.* **89**, 6055–6059 (1992).
- Labhart, P. & Reeder, R. H. *Nucleic Acids Res.* **13**, 8999–9009 (1985).
- Dunaway, M. & Dröge, P. *Nature* **341**, 657–659 (1989).
- Hsiang, Y.-H., Hertzberg, R., Hecht, S. & Liu, L. F. *J. biol. Chem.* **260**, 14873–14878 (1985).
- Paulson, J. R. & Laemmli, U. K. *Cell* **12**, 817–828 (1977).
- Mirkovitch, J., Mirault, M.-E. & Laemmli, U. K. *Cell* **39**, 223–232 (1984).
- Labhart, P. & Reeder, R. H. *Cell* **37**, 285–289 (1984).
- Pikaard, C. S., McStay, B., Schultz, M. C., Bell, S. P. & Reeder, R. H. *Genes Dev.* **3**, 1779–1788 (1989).
- Carroll, D., Wright, S. H., Wolff, R. K., Grzesiuk, E. & Maryon, E. B. *Molec. cell. Biol.* **6**, 2053–2061 (1986).
- Dunaway, M. *Genes Dev.* **3**, 1768–1778 (1989).

ACKNOWLEDGEMENTS. We thank J. C. Wang, N. Cozzarelli, R. Kanaar, L. Zechiedrich and members of N.C.'s laboratory for discussion, and J. Rine for support. This work was supported by a grant from the NSF (M.D.).

Study of allosteric communication between protomers by immunotagging

Janet E. Lindsley & James C. Wang

Department of Biochemistry and Molecular Biology, Harvard University, Cambridge, Massachusetts 02138, USA

LIGAND-INDUCED allosteric changes in proteins are important in their cellular functions and regulation^{1,2}, and both concerted and sequential examples are known³⁻⁵. The distinction has entailed elaborate analysis, however, and only a few systems have been unequivocally analysed. We have investigated the coupling between ATP usage and DNA transport by type II DNA topoisomerases^{6,7}, and one key question concerning allostery in these dyadic enzymes is whether ATP binding to one protomer can induce a concerted conformational change in the entire enzyme. Here we use an enzyme with one immunotagged subunit defective in ATP binding and one wild-type subunit to show that it can. Our approach should be generally applicable in the study of allostery and communication between members of a macromolecular assembly.

Binding of a non-hydrolysable β , γ -imido analogue of ATP (AMP-PNP) to yeast topoisomerase II induces a conformational change which switches one major SV8 endoprotease cleavage site from Glu 410 (site A) to Glu 680 (site B)⁶. A mutation G144I, changing Gly 144 to isoleucine, abolishes this conformational change, as well as the ATPase and DNA-transport activities of the enzyme. To examine whether AMP-PNP-binding in one subunit can induce conformational change in the other, immunotagged G144I and untagged wild-type polypeptides were synthesized in the same cell. The untagged polypeptide was always present in excess, as a fortuitous consequence of the expression vectors used for these polypeptides, and most of the tagged subunits are thus expected to be paired with untagged ones. In Fig. 1, tagged polypeptides are revealed by a tag-specific monoclonal antibody in an immunoblot of dimers digested with SV8.

With dimers of tagged wild-type subunits, addition of AMP-PNP switches the main cleavage site in the amino-terminal half

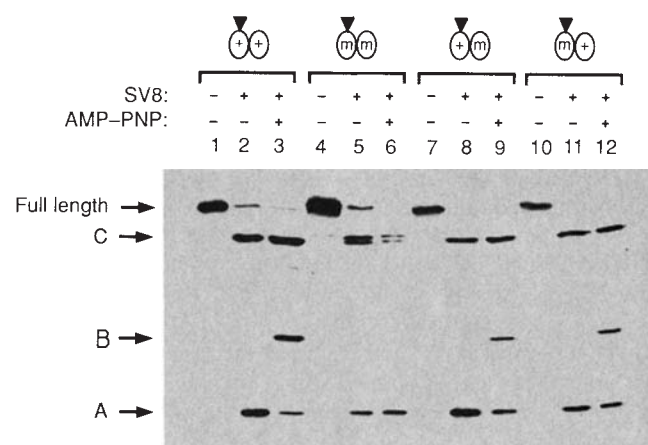
of the enzyme from A to B (lanes 1-3) as expected⁶; similarly, dimers of G144I subunits are cleaved mainly at A with or without AMP-PNP (lanes 4-6). Mixed dimers with either a tagged wild-type polypeptide (lanes 7-9) or a tagged mutant polypeptide (lanes 10-12), however, are cleaved mainly at A in the absence of AMP-PNP, but at B in its presence. The binding of AMP-PNP to the wild-type polypeptide thus appears to induce a concerted conformational change in both subunits, whether the second subunit is mutant or not. Because homodimers of the G144I mutant do not change conformation on binding AMP-PNP, the result in lane 12 also proves that heterodimers must form between mutant and wild-type polypeptides.

Recent experiments have led to the suggestion that type II DNA topoisomerases act as molecular clamps⁷. In the absence of ATP, the clamp is open and up to two DNA segments can enter its jaws. The binding of ATP closes the jaws, trapping the DNA. If two DNA segments are trapped, then one can be transported through a gate transiently opened by the enzyme in the other. When non-hydrolysable AMP-PNP is bound to the enzyme, the clamp is locked in the closed form, and any circular DNA within the clamp becomes topologically linked to the protein. Complexes between the enzyme, AMP-PNP, and circular DNA are thus stable in concentrated salt solutions^{7,8}, providing a further way to test for ligand-induced conformational changes.

Enzyme with an immunotagged G144I and a wild-type polypeptide was therefore first incubated with circular DNA; after further incubation with or without AMP-PNP, each mixture was analysed by CsCl density-gradient centrifugation. Without AMP-PNP, all DNA and tagged protein banded at positions expected for the free molecules (Fig. 2a). With AMP-PNP, however, a significant fraction of the tagged protein banded between free DNA and protein. From the relative molecular masses of the enzyme and DNA, the principal salt-stable species at 1.62 g ml⁻¹ consists of one dimeric enzyme per DNA ring, and the flanking bands correspond to species with one enzyme linked to two DNA rings and two enzymes linked to one DNA ring⁹. When both tagged and untagged polypeptides carry the G144I mutation, no salt-stable complex was detectable with or without AMP-PNP (Fig. 2b).

These experiments again indicate that the binding of ATP to one subunit of topoisomerase II can induce a concerted

FIG. 1 Immunostaining of partial SV8 endoproteolytic digests of purified yeast DNA topoisomerase II consisting of epitopically tagged and untagged polypeptides. A nitrocellulose blot of the digests following electrophoresis in an SDS-polyacrylamide (5-10% linear gradient) gel was probed with human *c-myc* monoclonal antibody MYC 1-9E10.2 as described⁶, except that the Amersham ECL system was used for detection. The symbol above each set of three lanes denotes the composition of the mixed dimer: + and m represent the wild-type and mutant polypeptide, respectively, and the triangular crown indicates the presence of the immunotag in the polypeptide directly underneath. In each reaction mixture, the purified yeast enzyme was pre-incubated for 10 min in the presence or absence of 1 mM AMP-PNP (as indicated at the top by a plus or minus sign). SV8 endoprotease was added as indicated, and partial digestion carried out as described⁶. A, B and C denote the main SV8 proteolytic products that retain the amino-terminal immunotag. Fragment C is formed by cutting at a strong site around amino acid 1,200; an additional cut at Glu 410 (site A) or Glu 680 (site B) yields fragment A or B, respectively⁶. The doublet nature of the C fragment in lanes 5 and 6 is probably due to incomplete cleavage within a cluster of proteolytic sites around C; these doublets were not seen in repeat experiments using different preparations of the various mixed dimeric enzymes. **METHODS.** The yeast strain JEL1 (α *leu2 trp 1 ura3-52 prb1-1122 pep4-3 Δ his3::PGAL10-GAL4*) was co-transformed with a pair of plasmids chosen from a set of four derivatives of YEPTOP2-PGAL1 or a homologue of it in which the *LEU2* marker replaces the *URA3* marker. The set of four plasmids overexpress tagged or untagged wild type or G144I mutant *S. cerevisiae* TOP2 gene (see ref. 6 for the construction of the URA3 marked plasmids; their *LEU2* marked homologues were constructed by routine methods).



Induction and purification of the overexpressed DNA topoisomerase II was as described⁶. Restriction analysis of plasmids rescued from the yeast cells used for protein purification showed no detectable recombination between the pair of plasmids they harbour, and thus there could be no transfer of the immunotag from the tagged polypeptide to the untagged one.

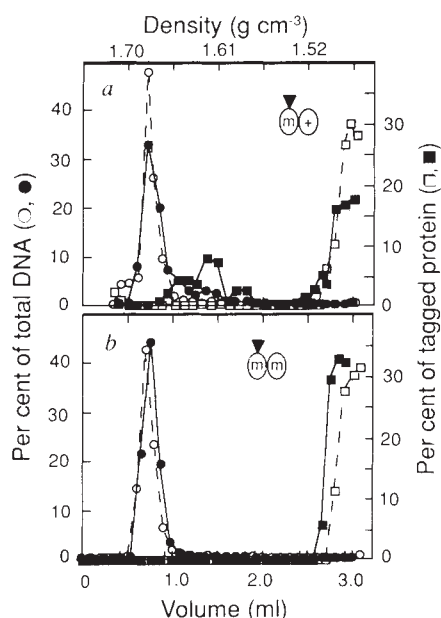


FIG. 2 Analysis of salt-stable complexes formed between yeast DNA topoisomerase II and DNA by equilibrium centrifugation in CsCl density gradients. The schematic illustrating the composition of mixed polypeptide dimers is described in the legend to Fig. 1. Curves with open symbols depict results with reactions without AMP-PNP, and those with filled symbols depict results with reactions containing AMP-PNP. The protein-to-DNA mass ratio in a salt-stable complex was estimated from its buoyant density as described⁹, assuming comparable hydration of protein and DNA. Each reaction mixture (1 ml) contained 50 µg of a 2.95-kilobase-pair ³²P-labelled relaxed plasmid DNA, about 35 µg yeast enzyme (containing several µg immunotagged polypeptide), and 50 mM Tris-acetate, pH 7.5, 50 mM potassium acetate, 8 mM magnesium acetate, 5 mM 2-mercaptoethanol, 10% (vol/vol) glycerol. Reactions were incubated for 5 min at 30 °C, and AMP-PNP was added to 1 mM final concentration to half of the reactions; incubation was continued for another 20 min, and 2 ml saturated CsCl containing 8 mM magnesium acetate was added to each of the reaction mixtures without AMP-PNP, and the same volume of the CsCl-magnesium acetate stock plus 1 mM AMP-PNP was added to each of the reaction mixtures with AMP-PNP. The final samples (3 ml) were placed in 5-ml polyallomer centrifuge tubes (Beckman), and each was overlaid with 2 ml mineral oil. Density gradient centrifugation of the samples was carried out at 38,000 r.p.m. and 20 °C in an SW55Ti rotor in a preparative ultracentrifuge (Beckman). After 3 days, fractions of ~100 µl each were collected from the bottom of each pierced tube. The volume indicated is the total volume from the bottom of the centrifuge tube. DNA was determined by scintillation counting, and the amount of the tagged protein by scanning immunostained dot blots of the fractions using a PDI scanner. Densities of fractions were estimated from refractive indices, without correcting for the presence of reagents other than CsCl and water.

conformational change in both subunits, and demonstrate the ease with which allosteric communication among components of a multimeric complex can be studied by immunotagging the mutant subunits in a mixed mutant and wild-type complex. □

Received 28 September; accepted 9 December 1992.

1. Monod, J., Wyman, J. & Changeux, J. P. *J. molec. Biol.* **12**, 88–118 (1965).
2. Koshland, D. E., Nemethy, G. & Filmer, D. *Biochemistry* **5**, 365–385 (1966).
3. Perutz, M. F. *Q. Rev. Biophys.* **22**, 139–237 (1989).
4. Schachman, H. K. *J. biol. Chem.* **263**, 18583–18586 (1988).
5. Fersht, A. *Enzyme Structure and Mechanism* 263–292 (Freeman, New York, 1984).
6. Lindsley, J. E. & Wang, J. C. *Proc. natn. Acad. Sci. U.S.A.* **88**, 10485–10489 (1991).
7. Roca, J. & Wang, J. C. *Cell* **71**, 833–840 (1992).
8. Osheroff, N. *J. biol. Chem.* **261**, 9944–9950 (1986).
9. Depew, R. E., Liu, L. F. & Wang, J. C. *J. biol. Chem.* **253**, 511–518 (1978).

ACKNOWLEDGEMENTS. We thank J. Roca for discussion and R. A. Kim for advice on density-gradient experiments. This work was supported by a grant from the US Public Health Service, and a postdoctoral fellowship from the American Cancer Society (to J.E.L.).

PLEASE follow these guidelines so that your manuscript may be handled expeditiously.

Nature is an international journal covering all the sciences. Contributors should therefore bear in mind those readers who work in other fields and those for whom English is a second language, and write clearly and simply, avoiding unnecessary technical terminology. Space in the journal is limited, making competition for publication severe. Brevity is highly valued. One printed page of *Nature*, without interruptions, contains about 1,300 words.

Manuscripts are selected for publication according to editorial assessment of their suitability and reports from independent referees. They can be sent to London or Washington and should be addressed to the Editor. Manuscripts may be dealt with in either office, depending on the subject matter, and will where necessary be sent between offices by overnight courier. High priority cannot be given to pre-submission enquiries; in urgent cases they can be made in the form of a one-page fax. All manuscripts are acknowledged on receipt but fewer than half are sent for review. Those that are not reviewed are returned as rapidly as possible so that they may be submitted elsewhere without delay. Contributors may suggest reviewers; limited requests for the exclusion of specific reviewers are usually heeded. Manuscripts are usually sent to two or three reviewers, chosen for their expertise rather than their geographical location. Manuscripts accepted for publication are typeset from the London office.

Nature requests authors to deposit sequence and crystallographic data in the databases that exist for this purpose, and to mention availability of these data.

Once a manuscript is accepted for publication, contributors will receive proofs in about 4 weeks. *Nature*'s staff will edit manuscripts with a view to brevity and clarity, so contributors should check proofs carefully. Manuscripts are generally published 2–3 weeks after receipt of corrected proofs. *Nature* does not exact page charges. Contributors receive a reprint order form with their proofs; reprint orders are processed after the manuscript is published and payment received.

Categories of paper

Review articles survey recent developments in a field. Most are commissioned but suggestions are welcome in the form of a one-page synopsis addressed to the Reviews Coordinator. Length is negotiable in advance.

Articles are research reports whose conclusions are of general interest and which are sufficiently rounded to be a substantial advance in understanding. They should not have more than 3,000 words of text (not including figure legends) or more than six display items and should not occupy more than five pages of *Nature*.

Articles start with a heading of 50–80 words written to advertise their content in general terms, to which editors will pay particular attention. The heading does not usually contain numbers, abbreviations or measurements. The introduction to the study is contained in the first two or three paragraphs of the article, which also briefly summarize its results and implications. Articles have fewer than 50 references and may contain a few short subheadings.

Letters are short reports of outstanding novel findings whose implications are general and important enough to be of interest to those outside the field. Letters should have 1,000 or fewer words of text and four or fewer display items. The first paragraph describes, in not more than 150 words and without the use of abbreviations, the background, rationale and chief conclusions of the study for the particular benefit of non-specialist readers. Letters do not have subheadings and should contain fewer than 30 references.

Commentary articles deal with issues in, or arising from, research that are also of interest to readers outside research. Some are commissioned but suggestions can be made to the Commentary Editor in the form of a one-page synopsis.

News and Views articles inform non-specialist readers about new scientific advances, sometimes in the form of a conference report. Most are commissioned but proposals can be made in advance to the News and Views Editor.

Scientific Correspondence is for discussion of topical scientific matters, including those published in *Nature*, and for miscellaneous contributions. Priority is given to letters of fewer than 500 words.

Preparation of manuscripts

All manuscripts should be typed, double-spaced, on one side of the paper only. An original and four copies are required, each accompanied by artwork. If photographs are included, five sets of originals are required; for line drawings, one set of originals and four good-quality photocopies are acceptable. Reference lists, figure legends and tables should all be on separate sheets, all of which should be double-spaced and numbered. Three copies of relevant manuscripts in press or submitted for publication elsewhere should be included with each copy of a submitted manuscript, and clearly marked as such. Five copies of revised and resubmitted manuscripts, labelled with their manuscript numbers are required, together with five copies of a letter detailing the changes made.

Titles are brief and simple. Active verbs, numerical values, abbreviations and punctuation are to be avoided. Titles should contain one or two key words for indexing purposes.

Artwork should be marked individually and clearly with the author's name and, when known, the manuscript number. Ideally, no figure should be larger than 28 by 22 cm. Figures with several parts are to be avoided and are permitted only if the parts are closely related, either experimentally or logically. Unlettered originals of photographs should be provided. Suggestions for cover illustrations, with captions and labelled with the manuscript number, are welcome. Original artwork is returned when a manuscript cannot be published.

Protein/nucleotide sequences should ideally be in the three-letter and not the single-letter code for amino acids. One column width of *Nature* can accommodate 20 amino acids or 60 base pairs.

Colour artwork. A charge of £500 per page is made as a contribution towards the cost of reproducing colour figures. Inability to pay these costs will not prevent publication of essential colour figures if the circumstances are explained. Proofs of colour artwork may be sent to contributors under separate cover from their galley proofs.

Figure legends should not exceed 300 words and ideally should be shorter. The figure is described first, then, briefly, the method. Reference to a method published elsewhere is preferable to a full description. Methods are not described in the text.

References are numbered sequentially as they appear in the text, followed by those in tables and finally by those in figure legends. Only papers published or in the press are numbered and included in the reference list. All other forms of reference should be cited in the text as a personal communication, manuscript submitted or in preparation. Text is not included in reference lists. References are abbreviated according to the *World List of Scientific Periodicals* (Butterworths, London, 1963–65). The first and last page numbers are included; reference to books should include publisher, place and date.

Abbreviations, symbols, units and Greek letters should be identified the first time they are used. Acronyms should be avoided whenever possible and, if used, defined. Footnotes are not used in the text.

Acknowledgements are brief and appear after the reference list; grant and contribution numbers are not allowed.

Supplementary information is material relevant to Articles or Letters which cannot, for lack of space, be published in full, but which is available from *Nature* on request.

Submission. Manuscripts can be sent to the Editor at 4 Little Essex Street, London WC2R 3LF, UK or at 1234 National Press Building, Washington, DC 20045, USA. Manuscripts or proofs sent by air courier to London should be declared as 'manuscripts' and 'value \$5' to prevent the imposition of import duty and value-added tax.

Magnetic bead capture of expressed sequences encoded within large genomic segments

Danilo A. Tagle, Manju Swaroop, Michael Lovett and Francis S. Collins

Magnetic bead capture utilizes biotin-streptavidin magnetic bead technology to isolate cDNAs rapidly from large genomic intervals, giving several thousand-fold enrichment of the selected cDNAs. The technique can allow parallel analysis of several large genomic segments of varying complexities and can be applied to the isolation of expressed sequences from various tissue sources.

RAPID identification and isolation of transcribed sequences from defined chromosomal regions is critical to the success of positionally cloning any gene responsible for an inherited disorder¹. Conventional methods of finding transcripts, such as using phylogenetically conserved sequences and CpG-rich fragments as hybridization probes, are laborious and cumbersome when the candidate interval spans several hundred thousand to a few million base pairs. Recently, several new approaches aimed at identifying and isolating transcripts from large genomic segments, such as from heteronuclear RNA derived from human-rodent hybrid cell lines^{2,3}, trapping of exons based on the presence of splice junctions^{4,5}, direct screening of cDNA libraries with total yeast artificial chromosomes (YACs)^{6,7} and enrichment of cDNAs using immobilized YACs^{8,9}, have been reported. Although each of these methodologies has certain advantages, each has limitations that dictate specific applicability. The isolation of human transcripts from somatic cell hybrids^{2,3} has the advantage of scanning entire chromosomes or subchromosomal regions present in the hybrids, but in general, only constitutively-expressed genes can be isolated from these cell lines. As exon trapping^{4,5} rescues transcriptional units directly from genomic clones, it does not directly rely on the availability of tissue- or temporal-specific cDNA libraries. However, thus far the method requires relatively small genomic inserts, such as cosmids, for trapping. In addition, cryptic splice sites in the human genome can cause false positives, and the technique will also miss intronless genes. Direct cDNA screening^{6,7} using YACs allows several hundred thousand base pairs to be scanned at once but because of the complexity of the probe, it suffers from a low signal-to-noise ratio, resulting in low reproducibility. In contrast, cDNA enrichment schemes^{8,9} are potentially powerful and rather straightforward. It is questionable, however, whether immobilizing YACs on membrane filters, as described in these original enrichment methods^{8,9}, provides easy ac-

cess for complex cDNA probes to find their targets.

More recently, our group^{10,11} and others¹² have modified the cDNA enrichment strategy by hybridizing biotinylated genomic DNA to cDNA probes in solution

and subsequently capturing the hybridized complexes using streptavidin-coated magnetic beads (Fig. 1). We have used this method¹⁰ to isolate cDNAs encoded over a 2.2-megabase (Mb) interval for the Huntington disease (HD) candidate re-

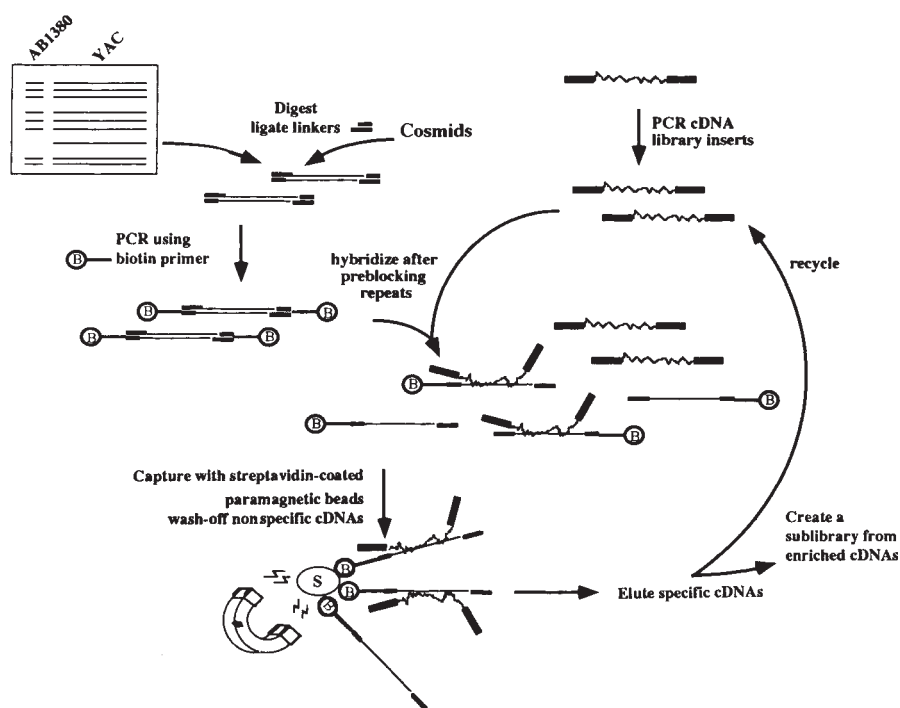


FIG. 1 General scheme for magnetic bead capture of expressed cDNAs from cosmids and YACs. Pools of miniprep cosmid DNA or of pulsed-field gel purified YAC genomic DNA are digested with four-base cutters and ligated to linkers. In place of restriction enzyme digestion, genomic DNA can also be sonicated and blunt-ended. The linked genomic segments are PCR-amplified using a 5'-biotinylated primer whose sequence matches one of the linker arms. cDNA library inserts are PCR-amplified using vector primers and hybridized at high stringency to the biotinylated genomic fragments after suppression of repetitive sequences. The biotinylated genomic-cDNA complexes are captured using streptavidin-coated paramagnetic beads and the unbound, nonspecific cDNAs are washed off. The captured cDNAs are eluted and PCR-amplified. The PCR products are digested with *EcoRI* and subcloned as a region-specific sublibrary. To increase selectivity, the captured cDNA material is recycled through a second enrichment. In a representative study, the minimal set of 8 overlapping YACs that spans the 2.2-Mb HD candidate interval from D4S126 to D4S98^{13,14} was used either individually or combined together as a pool. The YACs ranged in size from 200 to 600 kb and were deemed to contain only chromosome 4 sequences¹⁴. In addition, a pool of 17 arrayed cosmids (kindly provided by Gillian Bates of the Imperial Cancer Research Fund) that spans the region from D4S43 to D4S98 (approximately 550 kb) was used to capture cDNAs. Inserts from a commercially available random-primed, fetal brain cDNA library in λ -ZAP were amplified using T3 and T7 vector primers. A more detailed protocol is available from the authors upon request.

gion^{13,14}, which may contain as many as 100 genes in this very transcript-rich region of chromosome 4p16.3. Our comparisons using (1) pooled cosmid DNA, (2) individual YACs and (3) pooled YACs emphasize a key aspect of the methodology, that is, the technique is amenable to the parallel analysis of several large genomic segments of varying complexities in isolating expressed sequences from various tissue sources. As both genomic clones and cDNA probes are amplified by the polymerase chain reaction (PCR), the method provides another substantial advantage in that only a small quantity of starting materials is required. Specific 4p16.3 transcripts found in low abundance can be enriched several thousand fold with an associated decrease in nonspecific, ubiquitous transcripts, such as β -actin.

Technical considerations

Figure 1 shows schematically the general procedure involved in magnetic bead capture of cDNAs encoded from large genomic insert clones. The modification of biotinylating genomic DNA coupled with very high-affinity binding with streptavidin-coated paramagnetic beads provides better control of solution hybridization conditions compared to the kinetics of filter hybridization¹⁵⁻¹⁷. The strength and stability of the biotin-streptavidin coupling allows DNA manipulations, such as thermal denaturation and elution of annealed cDNAs or relative ease in changing buffers and wash solutions. This ensures a flexible system where preblocking, hybridization and washing can be performed easily. In addition, the beads are of equal size (monodispersed) and thus they follow uniform kinetics when subjected to a magnetic field¹⁸.

The biotin can be introduced into the genomic clones by photobiotinylation¹² or nick-translation¹² where, in both cases, biotin is incorporated randomly along the length of the DNA. An alternative way^{10,11} is to use a biotinylated primer during an *in vitro* PCR amplification reaction. The latter method has the advantage of tagging each amplified DNA strand at the 5' end with only one biotin molecule, which can prevent any possible steric hindrance with *Taq* polymerase during PCR or with cDNA probes during hybridization.

The genomic DNA used for capturing cDNAs can be from any genomic source. However in our experience, enrichment with cosmid DNA seems to work more efficiently than with YACs as during the gel-purification of the latter, YACs often co-migrate with degraded, higher molecular weight yeast chromosomes that harbour tandem ribosomal sequences. This has led to interesting enrichment artefacts where we^{10,11} and others⁸ often find that 30–60 per cent of the clones from the enriched sublibrary from YACs contain ribosomal sequences. This is most likely due to cross-species hybridization of highly conserved ribosomal

sequences between human and yeast and not from yeast contamination of the cDNA libraries, as the same phenomenon is observed when cDNA is extracted and used directly from cell lines¹¹.

A significant number of cDNA libraries contain repetitive sequences¹⁹ in the form of expressed repeats or unspliced messages and some cDNA clones contain repeats, particularly in their 3' untranslated region. Artefactually enriching these repeat-containing cDNAs can be partly avoided by suppressing repetitive sequences in the genomic target DNA with sheared human placental DNA. However, low-copy repeats are never adequately quenched, and can contribute up to 10 per cent of the enriched sublibrary. This is especially important as some chromosomal regions have relatively higher frequency of low-copy repeat sequences than others.

Specificity and enrichment

The specificity of the enrichment process can be assessed by examining the nature of the captured cDNAs. In the example of the HD region, PCR-amplified inserts from the starting fetal brain cDNA library were dot blotted along with amplified inserts from magnetic capture with each YAC from the HD candidate interval (Fig. 2). A cDNA probe, α -adducin, known to map telomeric to the D4S95 locus²⁰ which is spanned by two overlapping YACs (70D11 and YGA2) hybridized only to enriched cDNAs from

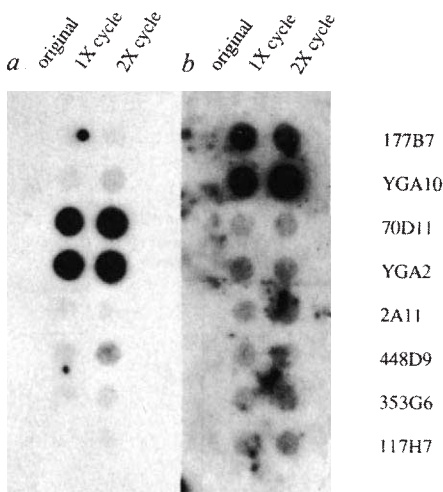


FIG. 2 Hybridization of control cDNA probes to captured cDNAs from YACs. The 8 YAC clones from the HD candidate interval¹⁴ used in the magnetic bead capture experiments were 177B7 (580 kb), YGA10 (450 kb), 70D11 (500 kb), YGA2 (600 kb), 2A11 (400 kb), 448D9 (350 kb), 353G6 (350 kb) and 117H7 (200 kb). Equal amounts (by weight) of PCR-amplified cDNA inserts from the starting library and from the once- and twice-enriched sublibraries for each of the YACs were dot-blotted onto membrane filter. Panels *a* and *b* show hybridization results when the blot was hybridized with the cDNA probes α -adducin²⁰ (encoded in the region spanned by YACs 70D11 and YGA2) and YC70²¹ (encoded in the region spanned by YACs 177B7 and YGA10), respectively.

these two YACs (Fig. 2a). Similarly, the YC70 cDNA probe only hybridizes to captured cDNAs using YACs 177B7 and YGA10 (Fig. 2b) which span across the D4S166 locus where YC70 has been mapped²¹. These demonstrate the high specificity of the captured cDNAs despite the complexity of the target DNA.

A quantitative assessment of the degree of enrichment can be determined by hybridizing cDNA probes to plaque lifts from original and twice-enriched cDNA libraries and counting the number of plaques detected by each probe. Approximately $1-2 \times 10^6$ and 1,000 clones from the original and twice-cycled cDNA libraries, respectively, were surveyed (data not shown). A screen of α -adducin²⁰ in the starting fetal brain cDNA library resulted in the detection of only one clone per 500,000 plaques. In comparison, α -adducin was represented in the twice-cycled cDNA library from pooled 4p16.3 YACs at a frequency of 1 in 250 clones, indicating a 2,000-fold enrichment of this cDNA. A cDNA clone, MBC6, isolated using pooled 4p16.3 cosmids was detectable in the primary cDNA library at 1 per 2×10^6 clones. MBC6 was detected in the twice-cycled cDNA sublibrary using the pooled 4p16.3 cosmids at 1 per 1000 plaques giving also a 2,000-fold enrichment. Similar levels of enrichment of cDNAs have been reported from a 425-kb YAC from chromosome 5 (ref. 11) and from two cosmid contigs (550 kb and 350 kb) from the X chromosome¹². As a corollary, we tested the level of nonspecific transcripts, such as β -actin, in the enriched material. β -actin was depleted at least 50-fold from the twice-cycled pooled YAC cDNA library.

To analyse the enriched cDNAs further, 50 clones from the twice-enriched cDNA sublibrary from the pooled 4p16.3 cosmids were randomly picked and mapped to a panel containing *Eco*RI digests of these cosmids (data not shown). Of the 50 clones, 35 mapped to at least one cosmid in the contig while the remaining 15 hybridized to cosmid vector bands, which indicates a possible contamination of the cDNA library with vector sequences.

Conclusions and prospects

Magnetic bead capture constitutes a powerful tool for the isolation of cDNAs encoded from large genomic regions. The technique is highly efficient and can be applied to many sources of cDNA and genomic clone pools in parallel. Furthermore, the method tends to normalize transcript levels, with a greater enrichment of rare messages than abundant ones. Several potential limitations, as well as possible solutions, are associated with this method. First, the resulting small insert sublibrary necessitates screening a full-length cDNA library to obtain full-length transcripts. The time involved in rescreening, however, can be cut substantially by screening random-primed cDNA

libraries. Alternatively, the insert size of the enriched cDNAs can be preserved by direct cloning of the PCR products. Second, genomic clones that contain region-specific low-copy repeats can cause a significant number of enriched cDNAs to contain these repeat sequences. Once identified, however, a running catalogue of these repeats can be used to pre-screen the sublibrary or can be incorporated as blocking agents in future magnetic bead capture experiments. Likewise, artefactually selected ribosomal cDNA clones can be pre-screened with ribosomal DNA probes. Third, since the methodology is based on hybridization by sequence homology, pseudogenes and members of multigene families may also be enriched. However, these can be sorted through by further mapping or sequencing of the enriched cDNAs. Last, in order to capture all encoded cDNAs from a given region, it would be ideal to use a cDNA library representing all possible transcribed human sequences. In this regard, it is possible to use pooled and normalized libraries from multiple tissue sources and from various developmental stages to approach such a complete cDNA library.

Magnetic bead capture of cDNAs provides an effective and straightforward alternative for isolation of expressed sequences from large genomic tracts. This method should be especially useful for isolating candidate genes where the defined candidate interval has been delineated by linkage analysis. It is anticipated that this technique will expedite the process of finding the

proverbial needle in a haystack for many disease loci. As the methodology is PCR-based, it may be also possible to extend the application of this technique to chromosome microdissected material, either as whole chromosome or as region-specific chromosomal scrapes, or to flow-sorted chromosomes, in order to generate a catalogue of chromosome-specific transcribed sequences that can be integrated with the genetic and physical maps of the genome. As such gene-hunting techniques become more efficient, isolating candidate genes will no longer be the rate-limiting step in positional cloning. □

Danilo A. Tagle¹ and Francis S. Collins^{1,2} are at the ¹Departments of Human Genetics and ²Internal Medicine, University of Michigan Medical Center, Ann Arbor, Michigan 48109-0650, USA. Manju Swaroop and FSC are with the Howard Hughes Medical Institute. Michael Lovett is at the Department of Biochemistry, University of Texas - Southwestern Medical Center, Dallas, Texas 75235, USA. DAT is supported by a National Institutes of Health - National Research Service Award and a Postdoctoral Fellowship Grant from The Wills Foundation. We thank our colleagues from the Hereditary Disease Foundation Collaborative Research Group, especially Gillian Bates and Hans Lehrach who have provided the ordered cosmids used in this study, Marcy MacDonald, Russell Snell and John Valdes who provided the control cDNA probes. We also thank Iris Gonzales for providing the human 45S rDNA probes, and Yossi Shiloh

and Lawrence Brody for helpful discussions. For more information on the protocol, fill in reader service number 100.

- Collins, F. S. *Nature Genet.* **1**, 3-6 (1992).
- Liu, P., Legerski, R. & Siciliano, M. J. *Science* **246**, 813-815 (1989).
- Corbo, L., Maley, J. A., Nelson, D. L. & Caskey, C. T. *Science* **249**, 652-655 (1990).
- Duyk, G. M., Kim, S. W., Myers, R. M. & Cox, D. R. *Proc. natn. Acad. Sci. U.S.A.* **87**, 8995-8999 (1990).
- Buckler, A. J. et al. *Proc. natn. Acad. Sci. U.S.A.* **88**, 4005-4009 (1991).
- Wallace, M. R. et al. *Science* **249**, 181-186 (1990).
- Elvin, P. et al. *Nucleic Acids Res.* **18**, 3913-3917 (1990).
- Lovett, M., Kere, J. & Hinton, L. M. *Proc. natn. Acad. Sci. U.S.A.* **88**, 9628-9632 (1991).
- Parimoo, S., Patanjali, S. R., Shukla, H., Chaplin, D. D. & Weissman, S. M. *Proc. natn. Acad. Sci. U.S.A.* **88**, 9623-9627 (1991).
- Tagle, D. A. et al. *Am. J. hum. Genet.* **51**, A22 (1992).
- Morgan, J. G., Dolgancov, G. M., Robbins, S. E., Hinton, L. M. & Lovett, M. *Nucleic Acids Res.* **20**, 5173-5179 (1992).
- Korn, B. et al. *Hum. molec. Genet.* **1**, 235-242 (1992).
- Bates, G. P. et al. *Am. J. hum. Genet.* **49**, 7-16 (1991).
- Bates, G. P. et al. *Nature Genet.* **1**, 180-187 (1992).
- Britten, R. J., Graham, D. E. & Neufeld, B. R. *Meth. Enzym.* **29**, 363-418 (1974).
- Britten, R. J. & Davidson, E. H. in *Nucleic Acid Hybridization - A Practical Approach* (eds Hames, B. D. & Higgins, S. J.) 3-15 (IRL, Oxford, 1985).
- Young, B. D. & Anderson, M. L. M. in *Nucleic Acid Hybridization - A Practical Approach* (eds Hames, B. D. & Higgins, S. J.) 47-71 (IRL, Oxford, 1985).
- Haukanes, B. -I. & Kvam, C. *Bio/technology* **11**, 60-63 (1993).
- Adams, M. D. et al. *Science* **355**, 632-634 (1992).
- Taylor, S. A. M. et al. *Nature Genet.* **2**, 223-227 (1992).
- Snell, R. G. et al. *Hum. molec. Genet.* (in the press).

On the biotech beat

Featured this week — an immunoglobulin binding protein, a chromosome enumeration system, a capillary electrophoresis system and columns, software, CD-ROMs and new product catalogues.

Two of the most thoroughly characterized members of the 'chaperonin' subgroup of molecular chaperones are the *Escherichia coli* proteins GroEL and GroES. Incubation of the inactive, denatured forms of several proteins, including carboxylase, citrate synthase, rhodanase, and a sigma subunit of bacterial RNA polymerase, with GroEL and GroES in the presence of Mg-ATP, results in *in vitro* reconstitution of their active forms (*Nature* **355**, 33-45; 1992). Epicentre Technologies now offers the **chaperonin proteins**, GroEL and GroES, together with another molecular chaperonin, DnaK protein, as tools for studying protein folding and assembly (*Reader Service No. 101*). The company says that the preparations of these proteins are rigorously tested for quality, both for the absence of contaminants and for functional activity. GroEL must demonstrate ATPase activity, and GroES, in 2.5-fold molar excess over GroEL, must

suppress greater than 90 per cent of the ATPase activity of GroEL. In addition, Epicentre Technologies says that its preparations of GroEL and GroES are tested for their ability to renature inactive citrate synthase in the presence of Mg-ATP to an enzymatically active form.

KappaLock is a new **immunoglobulin binding protein** from Aaston that binds to the kappa light chains of antibody types and, as such, should be useful for the purification of monoclonal antibodies and Fab fragments of antibodies (*Reader Service No. 102*). As binding is to the light chain at the Fab region of the antibody, the binding activity of the antibody is minimally affected. KappaLock is a 32-kD protein that is derived from a larger natural variant of a *Streptococcal* protein L. Natural protein L is a complex protein that binds albumin, cell wall components and lipids, in addition to

kappa light chains. Aaston says that the specificity of KappaLock has been improved over that of the naturally occurring protein L through the removal of the albumin binding and cell wall binding regions. It has four light chain binding regions and has been shown to interact with all subgroups of polyclonal antibodies from human, mouse, rabbit, pig and goat sources. As different antibody groups such as IgG, IgD and IgM have different numbers of kappa and lambda light chains, these different antibody groups can be expected to bind KappaLock with different strengths. This should allow for simple fractionation of antibody types with immobilized KappaLock.

The SpectrumCEP **direct-labelled chromosome enumeration system** from Imagenetics can be used for a wide range of applications in molecular cytogenetic research that uses fluorescence *in situ* hy-

ADVERTISEMENTS

Highest Quality Oligonucleotides

Synthesised, purified and shipped in 3 days

£2.75 per base

(All inclusive)

0.2 µM scale (approx. 500 µg)

- Antisense (completely or partially sulphurised)
- Extensive range of labelling and modifications
- Full design and analysis facilities
- Gene cassettes, bulk order and contract discounts

Molecular Medicine Unit

King's College School of Medicine & Dentistry

Tel. 071 - 326 3126 FAX 071 - 733 3877

Reader Service No.56

Synthetic Oligonucleotides

Rapid Quality Worldwide Service

£2.50 per base, no setup!

option of H.P.L.C. purification
biotinylation, fluorescent labelling
RNA Synthesis

phosphothioates, many other modifications

OSWEL DNA SERVICE

Department of Chemistry, University of Edinburgh
Kings Buildings, West Mains Road,
EDINBURGH EH9 3JJ

Tel: 031 650 4792 Fax: 031 662 4054

Reader Service No.15

Professional Custom Services

DNA SYNTHESIS
PROTEIN SEQUENCING
PEPTIDE SYNTHESIS

Rapid Service • Guaranteed Quality • Ample Quantity
Competitive Prices • Free Quotations
World-wide Delivery

V.G. Services

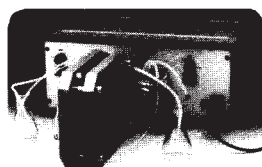
1200 Wonderland Rd. S. Bldg. #7 Unit #2
London, Ontario Canada N6L 1A8

CALL FAX: (519) 652-3758 ALT. FAX: (519) 652 0053

Reader Service No.32

LIPOSOMES

REPRODUCIBLE — HOMOGENEOUS —
EXTREMELY RAPID



Our new LIPOSOMAT guarantees rapid (5ml within 60 minutes) preparation of uniformly sized unilamellar liposomes. Liposome size is selectable between 25 and ca. 600 nm diameter. Sample volume is normally 5–10ml or up to 200ml if used in combination with an additional instrument. Lipid concentration can be up to 300 mg/ml.

DIANDROM

POB 126, D-8000 Munich 65, FRG

Reader Service No.23

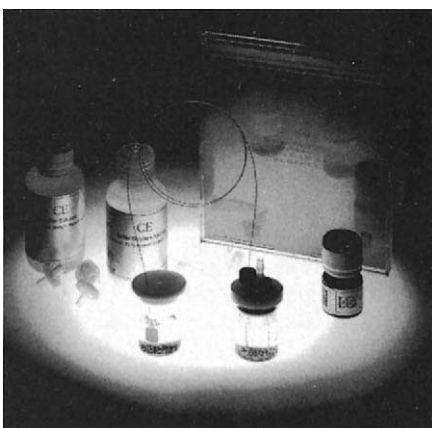
bridization procedures (*Reader Service No. 103*). With the SpectrumCEP, chromosomes can be counted in both metaphase and interphase cells. The system incorporates Spectrum-Orange and SpectrumGreen fluorophores, which are covalently linked to chromosome-specific satellite DNA probes. These direct-labelled DNA probes are designed to eliminate the potential for high background arising from the multi-layered detection reagents that are used in conventional, indirectly-labelled systems based on biotin/avidin or digoxigenin. In a single step, SpectrumCEP satellite DNA probes simultaneously hybridize and fluorescently tag each target sequence, assuring that hybridization signals are bright, compact and rapidly detected. The system comes complete with all the reagents necessary to hybridize and detect chromosomes in specimens that have been fixed to a conventional microscope slide. Test kits for performing ten assays include either SpectrumOrange- or SpectrumGreen-labelled satellite DNA probes, hybridization buffer, NP-40 detergent, propidium iodide and DAPI counter-stains, and SSC powder. Each reagent comes fully optimized for use with fluorescence *in situ* hybridization procedures, and a user manual contains several different assay options, including a standard four-hour hybridization protocol, the use of 'touch prep' procedures for solid tissue samples, modified ASG banding procedures, and methods for utilizing multiple probe hybridizations.

■ Capillary electrophoresis

The three new capillary electrophoresis columns from Isco can be used with most instruments and are designed to provide fast, high-resolution, reproducible separations for bioanalytical and pharmaceutical applications (*Reader Service No. 104*). Isco says that the CE-100/C18 capillary column significantly reduces problems caused by proteins sticking to the negatively charged wall of a conventional silica capillary electrophoresis column. CE-100/C18 uses a bonded C18 coating and surfactant buffer additive to produce a renewable, hydropho-

bic surface that minimizes sample loss and maximizes resolution. Electroosmotic flow is greatly reduced and essentially independent of pH, Isco says, improving the repeatability of migration times compared to untreated silica capillaries. The CE-200/glycerol column has an internal coating of glycerol bonded through a carbon chain linker to the silica surface. Isco says that compatibility with acetate and other nonsurfactant buffers makes this column suitable for capillary electrophoresis-mass spectrometry interfacing applications. By stabilizing electroosmotic flow via an anionic sulphonic acid coating, the CE-300/sulphonic column is designed to deliver more reproducible separations of small analyte molecules such as vitamins. It also provides a nonsurfactant alternative to MECC techniques for capillary electrophoresis of nucleotides, which can be separated using standard phosphate buffers.

Hewlett Packard hopes that with the introduction of its HP 3D capillary electrophoresis system, the transition of high-performance capillary electrophoresis technology from a research tool into a standard analytical technique will be accelerated (*Reader Service No. 105*). The fully automated system incorporates a new high-sensitivity diode-array detector designed for on-capillary detection and HP Extended Light Capillary Paths, which, the company says, increase sensitivity threefold over straight capillaries, without sacrificing resolution. The system is fully automated — from introduction of the sample to result reporting. The parameters that can be preset and performed at any time during the analysis include autosampling, multiple injection, preconditioning, off-line buffer replenishment, polarity switching, fraction collection and simultaneous programming of voltage and pressure. Full spectral data for compound identification and confirmation, and peak-purity assessment are available for each analysis. Hewlett Packard says that the peak-purity algorithm speeds up method development by determining the purity of nonGaussian peaks, which frequently occur in capillary electrophoresis. Random access is available to all 48 vials in the carousel for autosampling and fraction collection. Vials are sealed and the temperature is controlled thermostatically to prevent sample evaporation and degradation. Forced-air cooling controls capillary temperatures from 10 °C below ambient up to 60 °C. The system is controlled by an HPCE, 3D HP ChemStation. Although originally considered suitable primarily for the analysis of biological macromolecules, Hewlett Packard hopes that, with the introduction of its new system, capillary electrophoresis will be used for the separation of compounds such as amino acids, carbohydrates, chiral drugs, dyes, inorganic ions, oligonucleotides and DNA restriction fragments, organic acids, peptides and proteins, pesti-



New columns for capillary electrophoresis.

cides, surfactants, vitamins, and whole cells and virus particles.

■ Software and CD-ROMs

Release 6.7 of PC/GENE, the **protein and nucleic acid sequence analysis software** for IBM PCs and compatible computers developed by IntelliGenetics, now includes a new PCR primer design program called PCRPLAN and expanded capabilities in the existing programs for pattern searching (*Reader Service No. 106*). With PCRPLAN, users design a primer with as much assistance from the program as they desire, and then let the program evaluate it. If a proposed primer is rejected, the software explains why. PCRPLAN allows the user to control the location and length of the region to be amplified, in addition to melting temperature, GC content and length of the primer, and many other primer-binding parameters. Release 6.7 also includes enhancements to other programs used for pattern searching, RNA folding, physical chemistry and restriction mapping.

Biotech CAPS is the latest **biotechnology patent database** from MicroPatent, summarizing all the latest developments in the field and providing users with to a CD-ROM patent database of biotechnology claims and abstracts (*Reader Service No. 107*). The database provides industry research and development teams, technical librarians and patent specialists with a standalone research tool. Improved software allows the user to access the database of full claims and abstracts via 20 front-page search fields. Information can then be cross-referenced using 'windows', collated via a 'clipboard' function, tagged for future printing, or nested in order to allow searching at several levels. MicroPatent says that Biotech CAPS is published well ahead of most online services and cumulates to provide over 10 years of records on a single CD-ROM. The product is designed to complement the company's Biotech Patent-Images, which provides users with a regularly updated source of US biotechnology patent documents in facsimile format.

■ Reagent round-up

IBI's new DNA Quik Strip Kit can be used to determine the concentration of single- or double-stranded DNA, RNA or oligonucleotides as low as $10 \text{ ng } \mu\text{l}^{-1}$ in less than 3 min (*Reader Service No. 108*). It should be useful for **determining nucleic acid concentrations** from procedures such as library construction, plasmid/phage DNA preparations, cDNA synthesis, eluted nucleic acids, RNA transcription, polymerase chain reaction product monitoring, and poly(A⁺) mRNA isolation. The assay does not use radioactivity or ethidium bromide and provides permanent results without gel electrophoresis or photography. The procedure is as follows: the sample(s) is spotted on the Quik Strip, the strip is dipped in DNA

Blue Stain for 30 s, and then rinsed in water for 1–2 min. The concentration of the unknown sample is determined by matching the colour intensity of the unknown sample to standards provided in the manual.

Scotlab announces the introduction of the Nucleon **DNA extraction kit**, which is designed for the rapid and efficient isolation of DNA from complex samples such as whole blood or animal cell cultures (*Reader Service No. 109*). Use of the kit eliminates the need for phenol extractions and



For easy DNA extractions — see Scotlab's Nucleon kit.

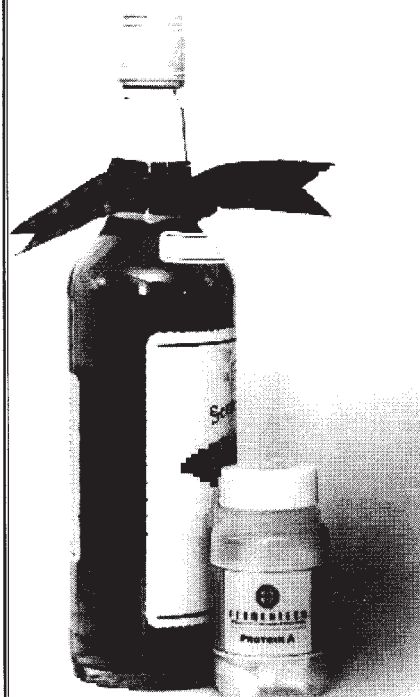
proteinase-K digestions. Scotlab says that DNA isolation takes less than 2 h and the extracted DNA is suitable for the polymerase chain reaction or restriction fragment length polymorphism applications. The kit can be used with both small and large sample volumes (5–30 ml). A specially designed microtube insert is designed to optimize the system for the recovery of small volume samples (less than 1.5 ml).

RNA phosphoramidites and columns are now available for synthesizing oligoribonucleotides on Applied Biosystems' Model 380B, 392 and 394 synthesizers (*Reader Service No. 110*). These new **RNA reagents** are used to produce oligoribonucleotides for ribozyme, protein binding, tRNA interaction, RNA antisense and other studies. They contain labile base-protecting groups that are removed in only 2 h at 55 °C. ABI says. This reduced deprotection time decreases the risk of internucleotide cleavage, phosphate migration and base modifications, resulting in an increase in both yield and chemical authenticity. Chemical authenticity of the full-length RNA oligonucleotide is essential for achieving biological or catalytic activity. Polystyrene, used as the support material rather than conventional controlled-pore glass, provides an inert, solid-phase support and increased hydrophobicity of the synthesis environment — two important factors in obtaining high-purity oligoribonucleotides. It also helps prevent extraneous chain growth off the support matrix, which results in a cleaner full-length product, ABI says.

ADVERTISEMENT

SCOTLAND

source of perfect Malt Whisky and pure Protein A



**both produced in the same
green way for perfection**

For great Scots quality and value

Protein A

Contact Laraine Aikman, MBA,
or Roy Fraser, B.Sc.,
or use reply card



FERMENTECH

medical limited
Heriot Watt University
Edinburgh EH14 4AP
SCOTLAND

Tel: UK 031 449 5055
International: 011 4431 449 5055
Facsimile: 031 449 7676

PRODUCT REVIEW

UniSyn Technologies has introduced a new kit for the **purification of goat IgG** (*Reader Service No. 111*). The kit is based on Avid AL, a synthetic affinity ligand that has been found to bind with high affinity to antibodies of many species whose immunoglobulins do not bind well to Protein A or Protein G. Avid AL is a nonprotein, non-carbohydrate compound with a molecular weight of less than 500 Da. As a synthetic ligand, Avid AL maintains functional activity after excessive exposure to heat, acid, base or protease treatment. These procedures, often used in the elimination of



Goat IgG affinity purification kit.

endotoxins from protein matrices, inactivate functional activity of Protein A or Protein G. When using Avid AL, binding and elution are carried out at pH 7.4, which circumvents the need for elution at acidic or denaturing conditions, the need for neutralization buffers, or the need for dialysis of the purified antibody. The purified goat IgG can be used immediately for conjugation purposes or for incorporation into an immunoassay using relevant reagents and buffer systems. Each millilitre of Avid AL gel has a binding capacity of 8–10 mg of goat IgG per millilitre of gel.

■ Cell transformation

BTX has developed a new **disposable cuvette**, each of which comes packaged with its own sterile cupette for easy specimen handling (*Reader Service No. 112*). The cupette is a specially designed dispos-

able Pasteur pipette with a narrow diameter, one-inch long tip for easy, complete removal of the electroporated cells. The BTX cuvette is moulded with embedded aluminium electrodes in 1-, 2- and 4-mm gap sizes. Each cuvette has a colour-coded, leak-resistant round cap that allows for quick and easy one-finger removal and easy specimen handling. The Cuvettes Plus are individually packaged and gamma-irradiated to ensure sterility and safety. BTX provides technical support and unlimited access to the Electronic Genetics Database.

For convenient subcloning, Invitrogen offers One Shot INVaF' **competent cells** (*Reader Service No. 113*). These cells are aliquoted in 50- μ l volumes. A transformation efficiency of 1×10^8 colonies per microgram of supercoiled vector is guaranteed, says Invitrogen. Bacterial strain INVaF' is designed as an all-purpose strain for library construction, single-strand rescue, and blue/white selection. One Shot cells are *recA*⁻ to reduce rearrangements and deletions and *endA*⁻ to eliminate internal endonuclease activity.

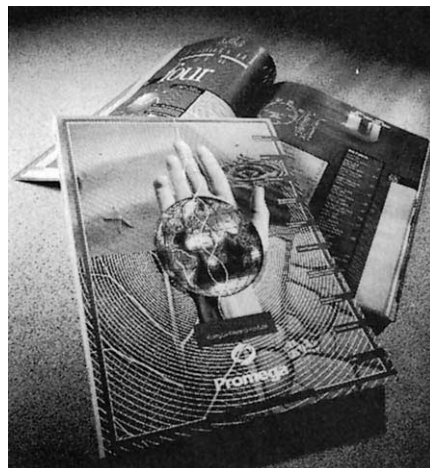
■ Shelf stockers

The latest edition of the ATCC **catalogue of plant viruses and antisera** can now be obtained from the American Type Culture Collection (*Reader Service No. 114*). The 84-page catalogue provides information on virus and viroid strains, polyclonal antisera and molecularly cloned viruses. Among the catalogue's sections are: an alphabetic listing of ATCC viroids, viruses and virus antisera strain descriptions; a list of special monoclonal antibodies; a taxonomic listing of viruses; an index to molecularly cloned viruses and viroids; and an alphabetic listing of plants susceptible to viruses in ATCC's collection. Copies are available free of charge in the United States and for US\$5.80 to all other foreign locations. The catalogue is also available on diskette, CD-ROM and via online access.

The identification and analysis of biochemical signal pathways is one of the main

topics in biomedical research. Biomol says that this is one of the reasons why the company has introduced a revised general catalogue for 1993 that focuses on analysis, fine chemicals and cellular regulation. In addition to the more established biochemicals, Biomol is offering **products for analysing the mechanisms of signal transduction**, including a variety of activators, inhibitors, ion channel modulators (*Reader Service No. 115*). The line-up includes a selection of signal transduction antibodies and proteins from Upstate Biotechnology — a US company for whom Biomol is the German distributor.

Promega's 1993 **catalogue** features many new products and more technical details about existing and new products (*Reader Service No. 116*). Some of the more novel products included are transcription/transla-



What's new from Promega for 1993.

tion coupled lysate/wheatgerm systems, magnetic particle-based purification systems for the isolation of mRNA, and more Magic products for the isolation of DNA. □

These notes are compiled by Diane Gershon from information provided by the manufacturers. For more information, fill in the reader service card bound inside the journal.

ADVERTISEMENTS



**Severn
Biotech Ltd.**

CUSTOM SERVICES

- **OLIGONUCLEOTIDE SYNTHESIS**
High quality PROBES, PRIMERS, ANTISENSE OLIGOS!
Fast-Efficient Service offering loads of modifications.
BIOTIN, FLUORESCENCE, 5' & 3' labelling.
Phosphothioates Affinity Support/Oligo's, etc. -
- **MULTIPLE PEPTIDE SYNTHESIS**
Epitope Mapping & Screening! We now supply 5mg quantities of high purity custom peptides up to 15 residues in length, using Fmoc chemistry. With D or L amino acids, these are supplied lyophilised in 4ml vials, cleaved from solid phase & ready for use.

Tel: 0562 825286 Fax: 0562 825284
Unit 23, Hootbrook Enterprise Centre
Worcester Road, Kidderminster DY9 1HY

Reader Service No.26

Synthetic

GENES

GGG GAA AAC GAA TCC
Gly Glu Asn Glu



**TIB
MOLBIOL**

for
further
information

FAX +49 30 7127734

Reader Service No.18

SINGER MSM SYSTEM

for YEAST GENETICS:-

- Ascus Dissection
- Pedigree Analysis
- Cell & Zygote Isolation

SINGER
INSTRUMENTS

"A responsibility to Science"

Call or Fax now for further information
& demonstration video

Tel: +44 984 40226 Fax: +44 984 41166

Singer Instrument Co Ltd, Roadwater, Watchet, Somerset, England, TA23 0QL

Reader Service No.13

US drug industry's research support

DRIVEN in large part by scientific and technological advances that make basic biomedical research a key element in drug discovery and design, many US pharmaceutical companies are increasing their funding of scientific research in universities and also at private research institutions.

"Industry realizes that the more classical way of discovering pharmaceuticals is not going to lead to the breakthrough drugs of tomorrow. That's why we're funding more research in disease mechanisms now," says Gary A. Wilson, director of science and technology at the Miles Inc. Pharmaceutical Division's research centre in West Haven, Connecticut. He predicts that his company's funding of external research will "grow by leaps and bounds" in the next few years.

As companies throughout the industry initiate and expand such funding efforts, a variety of programmes is emerging, often within a single company. Many programmes continue to evolve as administrators struggle with two dilemmas: how to manage money productively when it is sent outside the company and how to balance the company's scientific interests with those of external scientists.

Industry programmes that offer external funding for science fall into two principal categories. One kind of programme makes what are essentially charitable contributions to scientific institutions and scientists. Such programmes are usually administered through a non-profit fund set up by the company or by a group that administers other corporate charitable contributions, such as those to cultural and civic groups.

The other main type of programme is in support of scientific research entirely or chiefly because of its connection to the company's research interests and is often administered through the company's research division.

Debarred by tax laws from serving the donor company's specific interests, charitable programmes are usually more open than those linked to research divisions. They may fund numerous scientific disciplines, and often concentrate on significant but underfunded areas without apparent commercial potential.

Funding often takes the form of grants to upgrade research facilities or programmes. When individual assistance is offered, it usually goes to students or to young scientists in an effort to expand the pool of research talent. Because they are intended to serve needs rather than predetermined donor interests, such programmes are also more likely to hold open competitions for awards or to choose all grant recipients from proposals submitted through the mail.

One of the oldest and largest such programmes is the Burroughs Wellcome Foundation in Research Triangle Park, North Carolina, a non-profit private foundation that receives its funding from the Burroughs Wellcome Company. The bulk of its grants are awarded in open competitions in advance, many for young scientists. Executive director Martha G. Peck says the foundation continually tries to expand its applicant pool by targeting institutions and areas of the United States from which they have not received many applications in the past.

Many pharmaceutical companies have at least a small charitable programme, and most also fund basic scientific research and even scientific and medical education through programmes more

ing itself is for unrestricted use.

Such stipulations and others represent company attempts to balance the interests of the company, which needs a tangible return from its research investments, and those of the larger scientific community, which wants the funding to pursue its research questions in an atmosphere of intellectual freedom.

Some companies aim at balancing service to science with company research interests by using grants to form a strong community of distinguished researchers and institutional centres of excellence in particular disciplines. Examples include Alcon Laboratories Inc. of Fort Worth, Texas, which offers grants for eye research through its Alcon Research Institute, and Bristol-Myers Squibb Company, with headquarters in New York City, whose institutional research awards and career achievement awards cover seven different areas of research.

Grants in such programmes are normally for unrestricted use, and can be substantial in size. The programmes encourage or require recipients to remain in contact with each other and the company, by presenting their research at company-sponsored scientific symposia, nominating or selecting future grant recipients, or by other means.

For the relatively small number of scientists and institutions eligible for such awards, the attraction is the unrestricted and often generous funding. For the companies, it is the chance to act as benefactors to science by supporting research in a basic science area related to their own work, while ensuring maximum research value for their investment by limiting grants to individuals and institutions with proven track records.

Alongside their contracted research, which imposes restrictions and a bottom-line orientation many academic scientists may chafe against, many companies offer additional funding programmes especially designed to show researchers that industry does appreciate the values of the academic world.

At Pfizer Inc. of New York City, grants to young physician-scientists and funding of visiting professorships serve that purpose, according to Reeve Friedman, its manager of medical affairs. Before the company set up these programmes, in the early 1980s, it polled medical school deans to get the academic point of view about areas of need. Those polled "called physician-scientists an endangered species", so the Pfizer programme was set up to meet that perceived need, Friedman says. "I do get calls from PhDs who say, 'I'm mad. It's unfair to limit it this way.' But one company simply cannot do everything." ►

Recruitment and funding: special feature

This, the first of a series of bimonthly advertisement-linked features, is a survey of the ways in which pharmaceutical companies in the United States, Japan and Europe now support research and researchers from academic laboratories. The funds available from these sources seem to be growing rapidly, and are increasingly regarded by biomedical researchers as valuable supplements of the funds available from public grant-making agencies.

closely allied to their own research departments and frequently administered by them. It is in these programmes that the issues of managing money spent outside a company and balancing company interests against those of outside scientists become more acute.

Although the many research-driven programmes vary greatly, several features distinguish them from charitable programmes. Programmes more closely allied to company research interests often choose successful applicants from a limited pool of scientists, specifically targeted by company researchers or nominated by previous grant recipients.

In general, more of such a programme's funds are directed to scientists in mid-career and beyond, who have already achieved major successes, and to universities already on their way to becoming centres of excellence in some discipline. Recipients of grants from such programmes — even recipients of student scholarships — almost always enter some kind of contractual agreement with the company or are required to remain in contact with company scientists throughout the grant period and sometimes beyond, even when the fund-

Company	Amount of grants	Disciplines targeted	Grant types	How chosen	Contact address
FISONS CORPORATION	Various sized grants	Asthma, allergy	Grants to individual researchers or educational programmes relevant to company research interests	Company chooses from proposals submitted	Fisons Corporation, 755 Jefferson Road, PO Box 1710, Rochester, NY 14603
GENENTECH INC.	Various sized grants; approximately \$1,000,000 awarded in 1992	Medical, biological, molecular and genetic research, often related to company research interests	Scholarships, postdoctoral fellowships, endowed professorships, visiting scientist and visiting postdoctoral programmes, support for research projects, supply of recombinant protein reagents to academic investigators	Company scientists seek out recipients; some chosen from submitted proposals	Genentech Inc., 460 Point San Bruno Boulevard, South San Francisco, California 94080
HOECHST-ROUSSEL PHARMACEUTICALS INC.	A \$1-million 3½-year grant to Rutgers University from parent company, Hoechst Celanese Corporation — large portion of grant devoted to neuroscience on behalf of Hoechst-Roussel Pharmaceuticals Inc.	Neuroscience	Exploratory research grants, young faculty awards, graduate and undergraduate research awards	Joint company and university panels choose recipients	n.a.
JOHNSON & JOHNSON	\$2,500,000 in focused-giving grants awarded in 1992; 2-year to 4-year grants range from \$25,000 to \$100,000 a year	Basic research in health care and medicine, not necessarily related to company research interests	Grants to principal investigators in biomedical research at colleges, universities, and research institutes in the USA	Internal scientific committee chooses from submitted proposals	Johnson & Johnson Focused Giving Program, Johnson & Johnson, One Johnson & Johnson Plaza, New Brunswick, New Jersey 08933
MILES INC.	Grants of various sizes	Cardiovascular disease, microbiology, metabolic disorders (diabetes and obesity), dementia, bone and cartilage diseases (osteoarthritis and osteoporosis), inflammation, autoimmunity, synthetic chemistry, pharmacokinetics	Graduate fellowships, continuing medical education grants, research fellowships, continuing medical education grants, research institutions	Company scientists seek out recipients; some chosen from submitted proposals	Miles Inc., 400 Morgan Lane, West Haven, Connecticut 06516
PFIZER INC.	From \$6,000 visiting professorships through new faculty grants of \$65,000 a year for two years; approximately 30 visiting professorships and 12 postdoctoral and new faculty awards each year	Biological psychiatry/neuroscience, cardiovascular disease, diabetes/endocrinology, infectious diseases, rheumatology/immunology, clinical pharmacology, urology, geriatrics	All grants awarded to physician scientists (MDs and MD/PhDs); postdoctoral fellowships, new faculty grants, visiting professorships	Awardees chosen by independent peer review from submitted applications	Pfizer Inc., 235 East 42nd Street, NY 10017-5755
3M PHARMACEUTICALS	3M Foundation grants range from \$3,000 to \$15,000 — 2 to 3 grants run concurrently; other grants range from \$5,000 to \$100,000	Pharmacology (including immunology), toxicology (primarily mechanistic), medicinal chemistry	Support for research units; grants to support full-time junior faculty members at medical schools; career development fellowships for clinicians and PhD scientists; research starter grants, postdoctoral awards, undergraduate research awards, grants to institutions and individuals for purposes not included in formally announced competitive programmes	Chosen by PMAF Scientific Advisory Committee from submitted applications and proposals	Maurice Q. Bectel, President, Pharmaceutical Manufacturers Association Foundation, 1100 Fifteenth Street, NW, Washington, DC 20005
SYNTEX CORPORATION	Grants of various sizes	Chemistry, pharmacology, pharmaceuticals, cardiac research	Research awards, postdoctoral fellowships, other educational grants including scholarships and programme and operating funds	Company scientists seek out recipients of awards; some chosen from submitted proposals	Syntex Corporation, Corporate Contributions, 3401 Hillview Avenue, A1-194, Palo Alto, California 94304
* The leading US pharmaceutical companies were surveyed to establish main areas of interest and policy. The programs shown are not necessarily inclusive of the research sponsorship of the companies who chose to participate in this survey. The listings represent what each company considers its primary means of funding external scientific research.					

* The leading US pharmaceutical companies were surveyed to establish main areas of interest and policy. The responding companies are listed. The programs shown are not necessarily inclusive of the research sponsorship of the companies who chose to participate in this survey. The listings represent what each company considers its primary means of funding external scientific research.

Pfizer's grants for visiting professorships represent another way in which the company shows solidarity with academic values, according to Friedman. Academics already know all the ways in which corporate style differs from academic style, and the extent to which company programmes necessarily work to serve the company's interests, she says. "So if we are going to have a positive relationship with academia, we need to do some things to show that we won't always impose our opinions on them, that we respect values like intellectual freedom, too." Helping to bring visiting professors on to campuses is "one hundred per cent intellectually free," she says, "because the institution names who they want to come and makes arrangements with them. Then they ask, and the company gives them money to do it."

Other companies are trying other means of forging stronger ties with academic institutions, and sometimes with specifically chosen university scientists. Amgen Inc., of Thousand Oaks, California, for example, sponsors postdoctoral fellows in the laboratories of several scientists in the United States whose research is of interest to the company.

Amgen chose this traditional way of showing support for science in order to strengthen ties with researchers with whom it has previously worked and in the hope of "opening up avenues to researchers we'd like to have relationships with in the future," says Chris Giffin, associate manager of research affairs. "It's a way of opening up a dialogue with the principal investigator and the postdoc as well," Giffin adds.

Some research-division-based funding programmes offer scholarships to support undergraduate and graduate research, at nearby universities or at universities around the United States with strength in relevant scientific fields. Besides representing another goodwill gesture towards academic institutions, such scholarships are an attempt to attract well-trained students, with a positive view of working with industry, in to the research pipeline.

Gary Wilson, of the Miles Research Center, says the purpose of his company's graduate research programme is "to be a good neighbor, to encourage students to be researchers, and to hone the employee talent pool". As in other funding programmes sponsored by research divisions, recipients of Miles's graduate research funding are chosen by research centre staff, from a list provided by the university, on the basis of whether the students' research is of some interest to the company.

Later, the programme brings the students together with company scientists to

share the results of their work. "After about a year of research, we invite them to Miles to present their work," Wilson says. "They can tell a pretty good story by that time." The grants help to support the students' work, but, says Wilson, "it's certainly not all altruistic. If we have positions open, these are people we would like to consider."

In collaborative research projects closely tied to company goals, satisfying both the university and the company scientists becomes a special challenge, industry scientists say. Most companies are increasing their involvement in collaborative basic research with outside scientists, some small companies averaging one such collaboration for each scientist they employ. Contractual research agreements between the company and outside scientists vary widely from project to project. Contract terms range from offers of completely unrestricted funds to funding inextricably tied to product licensing agreements, with most projects falling in between.

Companies' research support can vary from offers simply to supply proprietary compounds or biological assets such as cell cultures, to multi-year support of \$100,000 or more. What does not vary, however, is the responsibility that research staff in industry have for achieving useful results from that funding, which can represent millions of dollars of a company's annual research budget.

As such contractual research programmes grow, the difficulties of effectively managing external research increases, company executives say. According to some, three features of their evolving collaborative research programmes are of particular importance: company scientists need to be proactive in seeking research to fund, to set clear goals for the research early in the project and to maintain close touch with the external scientists throughout the project.

Company scientists at Allergan Inc., in Irvine, California, are enthusiastic about Allergan's research collaborations, according to Larry Wheeler, vice president of biological sciences, and Mike Garst, vice president of chemical sciences. "The virtue for company scientists is that the company is not afraid to explore new things outside with academic groups," says Garst. "We have to fight for increased budgets to do this, but company scientists definitely want to do more." Wheeler and Garst say that Allergan's programme has gradually become more structured, and that, too, is satisfying to company scientists who feel responsibility for the way the money is spent. "We have learned from our experience that we have to be very goal-oriented," Wheeler says.

Although such severe goal-orientation may make it seem to some academic

researchers that for the company it is all get and no give, pharmaceutical industry scientists say that the companies bring some important elements to their scientific collaborations as well. Besides money, pharmaceutical companies can offer proprietary substances, mammalian cell cultures, monoclonal antibodies and other biological assets, as well as access to instruments and equipment such as X-ray crystallography facilities, which can make possible more adventurous research. In addition, says Allergan's Wheeler, academic investigators would like loyalty from their collaborators and funding agencies, "the knowledge that the funds are going to be there three or six months down the line, if that is what it takes to finish the project. You get that kind of loyalty through close interaction and pursuing the same goals."

To find academic collaborators whose goals intersect with those of company scientists, most pharmaceutical companies fund collaborative research projects throughout the United States and, in many cases, around the world. That circumstance can make it even more difficult to manage the money, as it is difficult for company scientists to make as many site visits as they might like. Hoechst Celanese Corporation, parent company of Hoechst-Roussel Pharmaceuticals Inc. of Somerville, New Jersey, recently initiated a pilot programme that takes a radical approach to that management issue.

Like other companies, Hoechst Celanese has previously offered its collaborative research funds to many institutions. In the pilot programme, the company will try to concentrate much of its scientific funding on one or two specific programmes at a single university campus near a major company facility (in the case of Hoechst-Roussel Pharmaceuticals, neuroscience programmes at Rutgers University). Joseph H. Patterson, executive vice-president of Hoechst Celanese, says the idea came from a principle at present gaining popularity in business: that narrowing the number of suppliers a company deals with makes the supply process easier to manage and increases the value of each relationship to both parties.

Each collaborative programme (initially, a \$1-million, 3½-year grant to each university involved) will provide funds to support exploratory research initiatives, support research by young faculty members and offer scholarships in relevant scientific fields for minority and female students as well as undergraduate and graduate research awards, and gradually build a variety of means — such as topical symposia and staff exchanges — whereby company scientists will interact with university staff and students.

Marcia Clemmitt

Japan's foundations still building

JAPANESE pharmaceutical companies, like their counterparts in the United States and Europe, are pouring increasing amounts of money into academic research. As in the West, a lot of the funding is charitable and goes into non-profit foundations that provide grants to researchers in Japan's universities and government research laboratories. But the process of selecting grant recipients is often less open compared with US and European foundations and tends to be controlled by powerful academics who head the grant selection committees.

Japan lacks the research-driven approach to funding by pharmaceutical companies often seen in the United States and Europe, but increasing numbers of Japanese pharmaceutical companies are establishing links with Western universities with long-term goals of drug development (see page 764). Another recent trend is a rush by Japanese pharmaceutical companies to fund endowed chairs at Japanese universities (see below), following a change in government regulations a few years ago that allows the establishment of such privately funded chairs.

There are dozens of non-profit research foundations in Japan supported by the pharmaceutical industry (see table overleaf for some examples). Most are

supported by individual companies but some, such as the Japan Health Sciences Foundation in Tokyo and the Senri Life Science Foundation in Osaka, are funded by groups of companies with national or local government support.

The top 25 foundations (in terms of funds) supported by individual companies distributed nearly ¥2,000 million in grants in 1991. More than half of the money was distributed as small research grants of ¥1 or ¥2 million (\$8,000–\$16,000), many of them for young researchers (under 40). About 10 per cent went to grants for overseas travel.

Nearly all the money is given to Japanese nationals but there is a growing tendency for foundations to provide support for non-Japanese researchers as well. The Takeda Science Foundation supported by Takeda Chemical Company, for example, provides more than 30 fellowships for medical doctors from South-East Asia to receive training in Japanese universities and research institutes. And Yamanouchi Pharmaceutical Company is planning to reorganize its foundation for research on metabolic disorders this year to establish an international section to fund non-Japanese researchers, says Teruhisa Noguchi, executive vice president of the company.

A few foundations openly advertise

their grant opportunities in magazines and elsewhere, but, in keeping with the Japanese tradition of relying on personal contacts, most depend on recommendations of candidates from powerful academics either within or outside the foundation. There is a growing tendency, particularly among the newer foundations, to use external peer review by committees of scientists to select grant recipients, but many foundations still use internal selection procedures. And "personal connections are still very important" in getting awards from most foundations, says Masanori Fukushima of Aichi Cancer Center, a vocal critic of Japan's medical research system.

Most Japanese companies do not expect direct returns on their investment in academic research. "Our foundation is 100 per cent charitable" says Noguchi of Yamanouchi, a view echoed by Tai Matsuzawa, managing director and vice president of the pharmaceutical group of Takeda Chemical Industries, who says "most of the foundations and research supported by Japanese pharmaceutical companies are set up without expecting a return . . . rather we aim at enhancement of the image of the company".

The Japan Health Sciences Foundation, which is supported by about 160 companies and the Ministry of Health

MESC's changes bear fruits in the universities

THE past three years have seen a rush by Japanese pharmaceutical companies to set up endowed chairs at Japanese universities following an easing of regulations by the Ministry of Education, Science and Culture in the late 1980s.

There are now about a dozen endowed chairs, supported with annual budgets of 30 to 50 million yen (\$250,000–\$420,000), and many more are expected to be established in the near future. Nearly all are in medical or pharmaceutical science faculties.

Takeda Chemical Industries, for example, is pumping 250 million yen (\$2 million) into Tokyo University's faculty of pharmaceutical sciences over five years to fund a chair in neuropathology and neurosciences. Selection of the appointment and running of the chair is done "totally" by the university, says Tai Matsuzawa, managing director and vice president of the pharmaceutical group of Takeda. The company may collaborate with the appointee by, for example, contributing research equipment, he says, but the main purpose is to "contribute to the promotion of education in pharmaceutical science". □

ENDOWED CHAIRS AT JAPANESE UNIVERSITIES

Company	University	Period of endowment	Total budget (million yen)
Tsumura & Co.	Gumma University	1988 to 1994	187
	Tokyo University	1991 to 1994	200
Sandoz Pharmaceuticals	Tokyo University	1991 to 1996	250
Eisai Co.	Tokyo University	1990 to 1995	150
Yokogawa Medical Systems	Shiga University of Medical Science	1989 to 1991*	60
Otsuka Pharmaceutical	University of Tokushima	1990 to 1993	130
Teikoku Pharmaceutical	Kagawa Medical School	1991 to 1996	100
Aso Pharmaceutical	Kumamoto University	1991 to 1996	n.a.
Takeda Chemical Industries	Tokyo University	1992 to 1997	250
Tanabe Seiyaku	Osaka University	1993 to 1998	300
Yamanouchi Pharmaceutical	Yamagata University	1993 to 1998	n.a.

* Since April 1990 has been extended each year with annual budget of 30 million yen.

TAKEDA FOUNDATION	YAMADA FOUNDATION	CIBA-GEIGY FOUNDATION
Foundation: Takeda Science Foundation	Foundation: Yamada Science Foundation	Foundation: Ciba-Geigy Foundation (Japan) for Promotion of Science
Supporting company: Takeda Chemical Industries	Supporting company: Rohto Pharmaceutical	Supporting company: Ciba-Geigy
Disciplines: Science, engineering, medicine, pharmacy, agriculture.	Disciplines: Natural science (physics, chemistry, medicine, biology).	Disciplines: Life science and chemistry/polymer science.
Type of grants: a) 50 research grants (total grant budget 60 million yen, maximum individual grant 10 million yen) for young Japanese researchers at 11 appointed universities. b) 33 grants (total budget 34 million yen) for non-Japanese medical doctors from Taiwan, Thailand, Korea, China, Philippines and Indonesia to train in Japan. c) Takeda prize for medical science. Two 3 million-yen-prizes for those who achieve significant contributions to medicine. d) Takeda Science Foundation Symposium on Bioscience is held in Kyoto every one or two years. Total budget for the 6th symposium (1990) 29 million yen.	Type of grants: a) 15 research grants of about 3 million yen each (total grant budget 55 million yen) for Japanese. b) 100 short-term (3-months) overseas travel grants of about 250,000 yen each (total budget 25 million yen) for Japanese. c) 10 long-term (6 months to 1 year) travel grants of about 1.5 million yen each for young Japanese. d) 20 short-term (3-months) grants for Japanese hosts to invite foreign scientists to Japan. Total grant budget 15 million yen.	Type of grants: a) 50 research grants of 2 million yen each for Japanese scientists in Japan. b) 10 Japan-Europe fellowships of 3.7 million yen each (covering travel and maintenance) for young Japanese scientists to do research in European institutions and for young Europeans to work in Japan. c) 10 grants of 0.5 million yen each to support international research meetings in Japan. d) Several travel grants of 0.2-0.5 million yen for Japanese scientists invited to give oral presentations at research meetings outside Japan.
Application/selection procedure: Recommendations for research grants are made by presidents and deans of the 11 appointed universities, by foundation counsellors and trustees, by members of the selection committee, and by the Pharmaceutical Society of Japan. Recommendations are screened by an internal committee and approved by the board of trustees. Grants for medical doctors approved by the board of trustees. Grants for medical doctors are recommended by selection committees consisting of members of medical societies and approved by the board of trustees. Prizes are recommended by people related to the foundation and former prizewinners, screened by an internal committee and approved by the board of trustees.	Application/selection procedure: The foundation solicits recommendations for candidates for research grants from university deans and directors of institutes. These applications are then screened by a committee and approved by the board of trustees. Applications for the other grants are made direct to the foundation and screened by the same process. No public appeal.	Application/selection procedure: Recommendations for grants for research meetings can only be made by trustees and counsellors of the foundation and selection is made by the board of trustees, the managing trustees, the chairman and vice-chairman of the Board of Councilors, and the chairman of the screening panel committee.
Contact: Takaaki Kamiya (executive trustee). Tel: 06-308-7418; Fax: 06-300-6034.	Contact: Kiyoshi Segami (standing director). Tel: 06-757-3311.	Contact: Ryo Sato (managing trustee) or Chuchi Takamashi (secretary general). Tel: 0797-74-2460; Fax: 0797-74-2325.

The Naïto Foundation		Eisai	
Fundamental research in natural science related to prevention and treatment of disease.			
a) 53 research grants of about 1.3 million yen each (total budget 68.9 million yen) for young researchers in above field.		a) 53 research grants of about 1.3 million yen each (total budget 68.9 million yen) for young researchers in above field.	
b) 22 grants of about 0.5 million yen each (total budget 10.6 million yen) for Japanese scientists who host foreign scientists in above field.		b) 22 grants of about 0.5 million yen each (total budget 10.6 million yen) for Japanese scientists who host foreign scientists in above field.	
c) 11 grants of about 1 to 3.5 million yen for overseas study by young (under 35) Japanese.		c) 11 grants of about 1 to 3.5 million yen for overseas study by young (under 35) Japanese.	
d) Five 3-year grants (maximum 3 million yen) for research in specific designated field of natural science that does not receive support from the government or other body.		d) Five 3-year grants (maximum 3 million yen) for research in specific designated field of natural science that does not receive support from the government or other body.	
e) 14 grants of about 0.5 million yen for travel to conferences and/or for research exchange.		e) 14 grants of about 0.5 million yen for travel to conferences and/or for research exchange.	
f) Five grants (total budget 2.5 million yen, maximum grant 1 million yen) for symposia.		f) Five grants (total budget 2.5 million yen, maximum grant 1 million yen) for symposia.	
g) Two grants (total grant budget 2 million yen) for publication of research results.		g) Two grants (total grant budget 2 million yen) for publication of research results.	
h) Five Naïto memorial grants (total grant budget 3.4 million) for urgent subsidy of research.		h) Five Naïto memorial grants (total grant budget 3.4 million) for urgent subsidy of research.	
i) Science promotion prize. 1 award per year for 3.0 million yen.		i) Science promotion prize. 1 award per year for 3.0 million yen.	
Recommendations of candidates for most grants are solicited from university deans or research institute directors or from trustees or counsellors of the foundation. Candidates for the science promotion prize, travel grants, symposium grants and the Naïto memorial grants are recommended by foundation counsellors and trustees. Candidates for grants for specific research (d) are recommended by a special committee. All grant applications are screened by an external committee and approved by the foundation councilors and trustees. All grants are for Japanese only.		Recommendations of candidates for most grants are solicited from university deans or research institute directors or from trustees or counsellors of the foundation. Candidates for the science promotion prize, travel grants, symposium grants and the Naïto memorial grants are recommended by foundation counsellors and trustees. Candidates for grants for specific research (d) are recommended by a special committee. All grant applications are screened by an external committee and approved by the foundation councilors and trustees. All grants are for Japanese only.	
Yasuo Kumagai (secretary general). Tel: 03-3813-3005.		Yasuo Kumagai (secretary general). Tel: 03-3813-3005.	

► and Welfare, on the other hand, does fund some research more directly related to drug development and is actively promoting joint research between government researchers and industry and technology transfer from the government to the private sector.

In 1991, the foundation distributed ¥1,413 million. Most of the money (¥1,090 million) consisted of small grants of a few million yen to more than 400 researchers in national institutes, universities and private companies. The rest went to research fellowships for Japanese and foreign scientists in Japanese national institutes, overseas travel grants, seminars and symposia, surveys, and joint research with institutes overseas. The foundation also supports a joint

programme to develop AIDS drugs, with a budget of ¥410 million in 1991.

There is, however, a darker side to some foundations funded by the pharmaceutical industry in Japan. It is common practice for powerful professors who have been involved in clinical trials to establish foundations on their retirement from university with donations from pharmaceutical companies. "Such foundations for clinical research are of questionable repute and help maintain the very poor level of clinical science in Japan", says Fukushima.

The pharmaceutical industry is, however, clearly trying to shed this dark image by investing more in basic research untied to drug development and sales.

David Swinbanks

Looking towards the West

LONG content to develop and sell drugs in Japan, Japanese pharmaceutical companies have only recently begun to look to universities and other centres of research excellence overseas to develop new ideas, a trend that is likely to expand rapidly in coming years. The long-term objective of the companies is to develop new drugs by tapping into the expertise of Western scientists. But the move by Japanese companies is also providing a welcome influx of cash to the universities.

One of the first to set up a research laboratory on foreign soil was Otsuka Pharmaceutical Company which opened the Biomembrane Institute in Seattle near the campus of the University of Washington in 1987. The institute, a non-profit cancer research organization, has close links with the university. Several of the staff are tenured members of the university, and graduate students and university researchers freely use the institute's facilities.

Yamanouchi Pharmaceutical has

adopted a similar approach in establishing the Yamanouchi Research Institute, UK next door to the science park of the University of Oxford in 1990. Teruhisa Noguchi, executive vice president of Yamanouchi, says the company chose the Oxford site because "Britain is very strong in molecular biology and cell biology, has many Nobel prizewinners, and has strong medical research organizations, such as the Medical Research Council".

Noguchi says he hopes to make a "strategic research alliance" between Yamanouchi and the university. The company is funding a postdoctoral research fellowship at the university's Wolfson College, and company researchers, mainly British, at the institute are collaborating closely with university departments and research units, including the Institute of Molecular Medicine. Yamanouchi has just reorganized its main research institute in Tsukuba science city northeast of Tokyo, and three new laboratories of molecular

medicine at Tsukuba will interact closely with the Oxford institute, says Noguchi.

Other Japanese companies have formed even closer links to Western universities by setting up laboratories on university campuses. Two examples are the Eisai London Research Laboratories Ltd at University College London, and the Fujisawa Institute of Neuroscience at the University of Edinburgh.

The Eisai laboratory is housed in the top two floors of a seven-storey new building built and equipped by Eisai at a cost of £12 million. The university has plans to occupy the other two floors, and the director of laboratory, Lee Rubin, who was recruited from the United States, says he has encouraged very active interaction with other researchers in the university.

The laboratory focuses on diseases of the nervous system, in particular cell death in neurons. And Rubin says "I think in five years time we will have made significant progress in achieving some understanding of what goes on in neurons when they die and ways of stopping them, and that 10 years from now we will be in the position to point in the direction of possible new products". Eisai plans to invest about £50 million in the laboratory over the next 15 years and "is in it for the long term" says Rubin.

In the United States, Noguchi says that Yamanouchi plans a "more flexible, amoeba-like approach" to links with universities and venture businesses, and he does not intend to establish fixed organizations such as the UK institute in Oxford.

Shionogi is providing a large grant (\$2 million over 6 years) to Nobel prizewinner Susumu Tonegawa at Massachusetts Institute of Technology for his research on immunology. Daichi Pharmaceutical has donated \$1.25 million to Vanderbilt University Medical School in the United States for an endowed chair in clinical pharmacology. And Ono Pharmaceutical Co. is pumping about £1 million a year into the William Harvey Institute at St Bartholomews Hospital in London to fund research in five areas related to inflammation of the arteries, tissue diseases (including rheumatoid conditions), cardiovascular diseases, the effect of hormones on inflammation and autoimmune diseases, and diabetes. The money supports 24 research staff and some support personnel at the institute.

The same company is donating £100,000 a year over five years to Kings College London to endow a chair in vascular biology, occupied by Jeremy Pearson. Pearson says this sum covers his salary and the costs of one research student. He says the company has put no restriction on his research and is keeping a relatively low profile.

David Swinbanks & David Dickson

LINKS BETWEEN JAPANESE PHARMACEUTICAL COMPANIES AND OVERSEAS UNIVERSITIES

Company	University	Type of link
Eisai	University College London, UK	On-campus laboratory
Fujisawa Pharmaceutical	University of Edinburgh, UK	On-campus institute
Yamanouchi Pharmaceutical	University of Oxford, UK	Near-campus laboratory
Otsuka Pharmaceutical	University of Washington, USA	Near-campus institute
Ono Pharmaceutical	1) Kings College, London, UK 2) William Harvey Institute, St Bartholomews Hospital, UK	Endowed chair Endowed research
Daichi Pharmaceutical	Vanderbilt University, USA	Endowed chair
Shionogi	Massachusetts Institute of Technology, USA	Endowed grant

The European pharmaceutical outlook

It is much harder to develop a new drug than it used to be. The costs of bringing a drug on to the market have soared: regulations have been tightened, competition is fiercer, and also — quite simply — the most obvious drug targets have already been exploited. To survive in the future, drug companies are having to seek out drug targets in less accessible, more complex biological systems, such as the brain and the immune system. This means more reliance on fundamental research, which cannot always be done cost-effectively in-house.

So drug companies are turning more and more to collaborations with academic institutes — giving badly needed money for research in exchange for rights to patent any interesting finding that may, in the long term, lead to a new drug. Nearly all companies have connections with universities and research institutes at some level. This can vary from funding a few lectures or PhD students to supporting major research projects or even whole institutes.

There are no hard and fast rules on reciprocal agreements, only various sets of guidelines. If the research work is done on a contractual basis, the company concerned is effectively paying for results. But, for collaborative research, no blanket policy exists; instead, an arbitrary agreement will be reached concerning results and intellectual property rights (IPR). Although the UK Committee of Vice-Chancellors and Principals formally recommends that "the IPR arrangements in European research contracts . . . state clearly that the intellectual property arising in a project is owned by the generating body", a sponsoring company will often negotiate a

deal to own intellectual property rights. But the agreement would allow the research group to use for its own purposes any knowledge gained. Should a collaboration lead to drug development, the academic body involved is likely to benefit from royalty payments. But it is difficult to put a cost to royalty payments because of the long and expensive processes of drug development and marketing.

In Europe, the activities of drug companies are diverse. The tradition of funding academic research in countries such as the United Kingdom, France, Germany, Sweden, Denmark and Switzerland, for example is particularly strong; most of the larger drug companies take part, and universities and research institutes the world over compete for their funding.

Two years ago, Glaxo Group Research in the United Kingdom announced that it would be putting more money into academic research. It was a move intended to "protect the company against too narrow a research focus" — a hope that widening its sphere of experience would throw up targets for tomorrow's drugs. Glaxo now supports collaborations with more than 80 per cent of British universities. Its major projects are with the department of pharmacology at the University of Bradford, the William Harvey Institute in London, the virology department at the University of Glasgow and the Institute for Rheumatic Diseases at Bath University. It also pays for chairs, senior lectureships and fellowships, and spends £2 million (\$3 million) a year on PhD studentships. It has recently made a £5 million investment in the department of pharmacology at the

University of Cambridge (see page 768). Glaxo says that its recent success means that it can afford to cast its net wider; as with other companies, it could be that it cannot afford not to.

The British-based Wellcome Trust is also a major supporter of academic research, spending globally £250 million on research and development (R&D), 30 per cent on basic research. Out of house, Wellcome sponsors a breast cancer research project at London's Royal Marsden Hospital to the tune of £1.8 million, and a £1 million clinical pharmacy project at King's College Medical School, also in London. The Wellcome Trust offers annually a few fellowships, 25 studentships and 50 sandwich course studentships as well as 'encouragement' awards in the form of prizes and grants. It also organizes about 70 symposia a year.

In the United Kingdom, most drug companies take part in projects of 'mutual scientific interest' with the Medical Research Council (MRC). The MRC tops up its limited research resources with around £5.5 million from the larger companies such as Pfizer, ICI, SmithKline Beecham, Wellcome and Glaxo as well as smaller ones such as Celltech and British Biotechnology. Collaborations are usually set up through the academic grapevine; researchers find industrial sponsors through the scientific network and vice versa. Many universities and research institutes are proactive in attracting industry's money. For example, the MRC set up a collaborative centre in 1986. This centre, with a staff of 62, identifies specific projects that might be attractive to industry and then markets them.

EUROPEAN PHARMACEUTICAL COMPANY PROFILES

Company	% Of turnover spent on R&D	% Of R&D spending on academic research	Studentships/fellowships	Symposia/conferences	Chairs	Institutes/university departments	Project grants	Prizes
Astra	18	15	✓		✓		✓	✓
Bayer	7	?	✓		✓	✓	✓	✓
Boehringer Ingelheim	15.5	1.5	✓	✓	✓	✓	✓	✓
Ciba-Geigy	10	?	✓	✓		✓	✓	
Fidia	25	?	✓		✓	✓	✓	
Glaxo	14	?	✓		✓	✓	✓	
Henri Beaufour	20	38	✓	✓		✓		
Hoechst	14.5	10						✓
Lundbeck	15.6	10.5	✓					✓
Novo Nordisk	13	1	✓	✓				✓
Organon	14.5	18					✓	✓
Pfizer	?	?						✓
Rhône-Poulenc	6.8	?	✓			✓	✓	
Roche	20	?	✓			✓	✓	
Roussel Uclaf	?	?	✓	✓			✓	
Sandoz	14	?	✓			✓	✓	✓
Schering	15	?	✓	✓	✓	✓		✓
SmithKline Beecham	9.1	?						
Solvay Duphar	13	?	✓	✓				✓
Synthelabo	20	?						
Wellcome	14	30	✓	✓	✓	✓	✓	✓

In Germany, academic research is supported by most of the major companies. For example, Bayer, one of the biggest investors in research and development worldwide, spent 7 per cent of turnover — or DM3 billion (\$1.8 billion) — on R&D in 1992. Nearly 40 per cent of this budget is spent on health research; two thirds is spent in Germany. It sponsors studentships and fellowships, a chair at the Institute of Genetics in Cologne and a range of prizes including the DM60,000 Otto Bayer prize. Bayer also supplies medical students with academic literature for research, a popular service that attracts between 100 and 150 written requests daily. Bayer also supports a DM4 million multicentre project in genetic engineering, related to its own drug development programme.

The pharmaceutical industry is regarded with some suspicion in Germany, and companies admit that their support of research in academic institutions is intended to improve the company's image by returning, without strings, part of their profits to academics. With an annual fund of around DM5 million, the Boehringer Ingelheim Fund targets young scientists with studentships and fellowships in fundamental medical research. Since its inauguration in 1983, the fund has awarded 336 grants and at present has almost 150 fellowships, around a quarter of the number offered by Germany's research council, the *Deutsches Forschungsgemeinschaft*. In 1992, Boehringer Ingelheim gave an additional DM5.8 million to research, including two posts at the University of Mainz in clinical and theoretical medicine. The company has established prizes and scholarships in Italy, Spain, the United States, Canada and the Philippines. It also organizes the international Titisee conferences, which deal with current topics in basic biology.

Academic research support from Schering AG Berlin (not to be confused with the US company Schering Plough) is also steered by a research foundation. It gives 12 fellowships/studentships, funds a chair in developmental molecular biology and biochemistry at the Free University of Berlin and has endowed the Institute of Organic Chemistry at the Technical University of Berlin. The company invests annually DM100,000 in prizes for basic research and endocrinology. With Astra-Medica in Frankfurt and Ciba-Geigy in Basel, the company sponsors a cancer research project at the Tumour Biology Clinic in Freiburg.

Within Scandinavia, the pharmaceutical industries of Sweden and Denmark are the most supportive of academic research. For example, the Swedish Astra group spends 18 per cent of turnover on R&D, although up to 85 per cent of this is spent on clinical trials. Astra

Fidia — hands across the ocean

FIDIA is a name that all pharmacologists and neuroscientists know. For the past several years it has popped up in association with virtually all areas of scientific activity. Lavishly sponsored meetings, authoritative books, seemingly limitless research funds. But, from the outside, it is also becoming apparent that the high-profile money is now drying up. A tightening of belts at the company's Italian headquarters in Padua has clipped the wings of the Fidra Research Foundation, set up in 1985 to fund cultural activities relating to science. But one project that has not been touched is the Fidra-Georgetown Institute of Neurosciences founded with Georgetown University in Washington DC. This institute is expected to move into its new building next year, and has another 20 years of funding guaranteed.

The institute began life in 1985 as a group of 20 scientists setting up shop in close collaboration with Georgetown University. It now has 75 staff. They are paid by a grant from the Fidra Research Foundation to Georgetown University, which, for an 18 per cent overhead fee, administers all the appointments and research finances. All staff hold joint appointments with appropriate university departments. Their appointments are recommended by the institute but selected by the department, which has complete control over their promotion. Each member of staff must spend 20 per cent of his or her time working for the department (teaching and so on), but the rest of the time is spent on research in institute laboratories. As members of the university, staff members apply for grant money to sources such as the National Institutes of Health (NIH), although Fidra pays for much of the equipment

and materials.

Although the planned new building was originally intended for the institute alone, the university asked to share it. This agreement, which delayed building for a few years, means that all Georgetown's neuroscience activity (70 per cent of which is done by Fidra) will be housed together when the building is completed next year.

Most of the work at the institute centres on GABA and glutamate neurotransmitter systems (from molecular biology to behavioural approaches) and therefore addresses concepts such as synaptic plasticity, anxiety and neurodegenerative disorders. Fidra has first rights on any patentable item arising from work at the institute, but the agreement between the two parties gives the university a 40 per cent share of royalties. There are two products now approaching the development stage. An antipanic agent is going through toxicity testing prior to Phase 1 trials, and a proposed antiepileptic is already in Phase 1. Both are benzodiazepine-like structures.

In addition to supporting the institute, the Fidra Foundation continues, on a modest scale compared with earlier times, to sponsor lectures, publications and meetings. A change in management at the parent company coinciding with a new foundation president resulted in a reduction of funding for these activities and a halving of foundation staff. The news bulletin *Neuroscience Facts* continues to be published from the institute twice monthly and is sent without charge to 6,000 scientists throughout the world, including 3,500 North American scientists who receive it by fax. This in itself keeps Fidra's public profile high.

Alison Abbott

firmly believes in decentralization of research; it has set up research centres in regions with established research environments such as Stockholm, Lund and Gothenburg, at universities, research institutes and teaching hospitals. It also funds 35 adjunct professorships; holders of these posts split their time between company and university work. Collaboration with the academic world, so the company believes, is most fertile when done informally: 'small is beautiful'. Astra's head of research Claes Wilhelmson says this policy "will continue to provide many ideas for future drugs".

In Denmark, research spending by the large drug companies is in the hands both of the companies and their related foundations. The foundations, set up

primarily to prevent unfriendly takeovers, commit an amount for academic research spending. This contribution remains constant, whereas the R&D cost varies according to annual sales.

Novo Nordisk maintains control of research policy by issuing two types of shares. 'A' shares are the sole property of the foundation and entitle the holder to majority voting rights: all of these shares are held by the Novo Nordisk Foundation. This gives the original founders of the company control of policy decisions, including those on R&D, and allows them to resist pressure from public shareholders.

Novo Nordisk spends 13 per cent of its turnover on R&D. Of this, Dkr14.5 million (\$2.3 million) is spent on academic grants, Dkr2.5 million on sym-

posia and DKr250,000 on the annual Novo Nordisk prize. Novo Nordisk is the world's largest manufacturer of insulin and owns the Steno Diabetes Centre in Gentofte in Denmark, which carries out research as well as offering treatment.

The Lundbeck Corporation, a private drug company owned by the Lundbeck Foundation, is the other major player in industry-sponsored academic research in Denmark. The foundation granted DKr16 million in 1991 for scientific and medical research and offers the annual DKr300,000 P. V. Petersen prize for cancer research. Poul Krosgaard-Larsen, of the Royal Danish School of Pharmacy, is very positive about industry's support of academic institutions. He says that collaborative projects can be set up quickly and easily because of the relatively tight-knit scientific community in Denmark.

In the Netherlands, both academics and industrialists are resentful of a negative government attitude towards the pharmaceutical industry and related research. The government says that it does not want to fund further an industry that it considers is too rich anyway. "Drugs are prices and prices should be cut", or so Koen Wiedhaup of Organon interprets government policy of creating a limited list of drugs and drug prices. He says that the large companies can still contribute to academic projects but further collaborations, vital to innovative research, are being hampered.

Dutch academics must continue to look to companies in other countries for more support. They have formed a lobby to challenge relevant government departments on their policy. This campaign, led by Dick Maier, has stirred the government to set up an independent investigation into the state of pharmaceutical research in the Netherlands.

Dutch researchers have adopted a policy of self-help. Henk Timmermann of the Free University of Amsterdam is trying to attract money from Japanese companies. Eisai and Yamanouchi are already involved in collaborative projects with the Universities of London and Oxford in the United Kingdom (see page 768). The University of Leiden and the Free University of Amsterdam have formed a consortium, with the University of Uppsala in Sweden and the University of London, which makes up for the lack of support for pharmaceutical sciences on a national level. The member universities set up staff exchanges and have a programme of conferences such as the Ulla summer symposium 1993.

The two major drug companies supporting academic research in the Netherlands are Organon and Solvay Duphar. The latter spends Dfl120 million (\$65 million) a year on R&D, 30 per cent

of which is for drug development. But out-of-house spending is restricted to Dfl2.0-Dfl2.5 million for postdoctoral fellowships and sponsorship of the Ariens lecture.

Organon spends 14.5 per cent of turnover on R&D, 18 per cent of which is out of house and shared equally between clinical and non-clinical research. Academic research collaborations in Europe are concentrated in the Netherlands and the United Kingdom. Organon supports programmes in endocrinology and pharmacology and is to introduce an annual prize for fundamental research related to drug development which will be overseen by the Dutch Royal Academy of Science. Organon also offers the Dfl150,000 annual family planning award related to contraceptive practises and to their drug Marvelon, the most widely prescribed oral contraceptive.

Both government and industrial support of academic research in Italy is dismal and is unlikely to improve as the country moves deeper into recession. But industry offers one remarkable exception in the shape of Fidia (see page 766). This Italian pharmaceutical company spends 25 per cent — or L20 billion (\$13 million) — of annual turnover on R&D worldwide. Fidia is the unrivalled star in Italy when it comes to sponsoring fundamental research. Italian companies, among them Sigmadau, Technofarmacia, Dombe, Pharmitalia and Letitet, restrict their academic research spending to buying expertise for applied research projects. Ironically, Sigmadau offered a fellowship to the Institute of Pharmacology at the University of Palma but it had to be turned down because the university could not find the required expertise. This shows Italy's problem — research scientists are voting with their feet and leaving behind cash-strapped university departments and institutes.

Switzerland has a great tradition of pharmaceutical companies; the big three — Hoffmann La Roche, Sandoz and Ciba-Geigy — are among the top 10 pharmaceutical companies in Europe.

Sandoz has an annual research budget of SFr1.4 billion (\$920 million). The company, commonly considered the slowest and most inward looking of the big three Basel-based companies, is moving into a new era of internationalization. It recently announced a US\$3-million agreement with the largest biomedical research institute in the United States, the Scripps Research Institute in California. Sandoz has also invested US\$100 million in oncology research at the Dana Farber Cancer Institute in Boston and a further US\$75 million in the Neurosciences Institute in New York. This spectacular spending, spreading money worldwide for long-

term research projects, represents a change in policy for the company. Sandoz previously relied largely on in-house research, but group director Marc Moret recognized that the company had to invest at the cutting edge of academic research. Even though management might lose precise control of research projects, he says, free and innovative research is the way in which pharmaceutical companies now enhance their long-term chances in an increasingly international market. The company spends SFr10 million on research into ageing, including the annually awarded Sandoz prize.

The largest Swiss pharmaceutical/chemical concern, Ciba-Geigy, spent 10 per cent of its 1991 turnover on R&D. In 1970, Sandoz set up the Friedrich Miescher Institute (FMI) as an independent foundation. The FMI's 120 scientists concentrate on cell and molecular biology and neurobiology. The FMI's contingent of scientists is international and the institute pays full travel and living expenses, although the study grant itself is not covered. In 1990, Ciba-Geigy agreed to give US\$20 million for arthritis research at the University of California and established an educational trust fund, the SFr25 million Ciba-Geigy Jubilee Foundation. The foundation puts further education grants at the disposal of graduates in Switzerland for scientific and medical research. Ciba-Geigy spread its wings still further with its support for the Foundation for the Promotion of Science in Tokyo, set up in 1987; it is the first foundation run by a non-Japanese company to receive authorization from the Japanese Ministry of Education, Science and Culture. Its objectives are to promote creative research in Japan and to exchange scientists between Japan and Europe.

Hoffmann La Roche spends a remarkable 20 per cent of sales on R&D, a figure reckoned to be the highest in the pharmaceutical industry. Roche's research programme creates an international picture. The company has a long-term R&D agreement with the Harvard Medical School, and with the Genetics Institute on AIDS Research. It also has the first centre for basic biological research to be financed by the pharmaceutical industry, the Roche Institute of Molecular Biology in New Jersey. This institute has around a hundred scientists. The Roche African Research Foundation in Abidjan tests about 20 Roche drugs for activity in infectious diseases and funds epidemiological research as well as research into AIDS, malaria and other infectious diseases. The foundation receives grant requests from researchers from all over West Africa. In Switzerland, the Roche Research Foundation gives grants to candi-

dates who are chosen from Swiss academic institutions according to the specific requirements of the foundation. The Basel Institute for Immunology, which houses 50 basic-research scientists, is also financed by Roche.

There are a number of reasons why Swiss drug companies invest internationally. First, expertise cannot always be found nationally, even when a company has vast in-house R&D operations. Worldwide collaborations also reflect Swiss companies' sales; for example, 96 per cent of Sandoz sales are outside Switzerland. Drug companies are discouraged from new long-term investment in Switzerland because of strict laws on animal experiments. A recent referendum rejected a complete ban, but drug companies remain wary.

France's largest pharmaceutical company, Rhône-Poulenc, sponsors much academic research. But this may be about to change. The government is at present selling 20 per cent of the shares to the public, but the conservative opposition, tipped to take over power after the elections later this year, has pledged to privatize the company completely. This move could ultimately

affect non-commercial academic research projects. Rhône-Poulenc's academic research activity is guided by the Directions Scientifiques. The company offers studentships and also, to the country's best young scientists, the chance of doing strategic research for the company instead of military service. Rhône-Poulenc partly sponsors the Bioaviner programme with the Ministries of Research and Industry. The project costs FF1.6 billion (\$290 million) and aims to define the basic research needs of the company. Rhône-Poulenc also spends FF0.2 billion on a five-year project with public research institutes. The institutes set the scientific agenda.

Roussel Uclaf has seven research agreements with universities and research institutes in France, including the state-financed CNRS, INSERM and INRA. The company has similar agreements with European universities and also attracts scientists for the Tables Rondes, a forum for scientific debate independent of company business. The biennial Roussel prize is awarded to a scientist researching in chemistry or biology of steroids. □

arrangement with Squibb means that approval for any paper arising from funded work (but not work funded from other sources) has to be sought. In practice this means a delay of around a week, unless there is material that may be of patentable value, in which case a delay of six months can be requested. But because Squibb is happy to consider a draft with just a summary of the results, says Smith, there have so far been no delays at all.

At about the same time, Cambridge was approaching industrial collaboration in a different way. When the physically scattered department of pharmacology was offered the opportunity to move to a new location on campus, of 70,000 square feet, department head Alan Cuthbert raised finance from independent sources to build new laboratories. The building included a suite of laboratories to be hired out on a long-term basis to a drug company, on "mutually beneficial terms". It was three years before Glaxo, which had earlier been approached, entered a ten-year agreement to move a team of research scientists into the laboratory space, now known as the Glaxo Institute of Applied Pharmacology (GIAP). As well as funding its own research, Glaxo has given £5 million to support research within the department. Its total investment over the ten years is £16 million.

There are important differences between the Oxford and Cambridge arrangements, chiefly that company scientists work directly with academics at Cambridge (there are no Squibb personnel in Oxford) and that Cambridge can spend its money entirely as it chooses. There are no restrictions on research topic and decisions about allocation of money are taken within the department. On the other hand, as at Oxford, funded scientists still have to let Glaxo consider their papers before publication. Glaxo has three weeks to turn them around, unless there are items of potential patent interest.

The Oxford and Cambridge experiences has so far proved very happy, in spite of early suspicion borne of historical distrust by academics of industry and its commercial interests. Times have changed and compromises have been made on both sides. For the universities, the injection of secure funds is an immediate reward, and the carefully worded contracts mean that little academic freedom is lost. For Squibb and Glaxo, the risk — though calculated — is greater. It will clearly be at least a decade before the true value to industry of such ventures will be established. But the willingness of these companies to make long-term investments reduces the risk of failure.

Alison Abbott

Making partnerships work

In the 1980s, the departments of pharmacology at the Universities of Oxford and Cambridge found themselves in a similar position. Like everyone else, they faced a future of diminishing research funding opportunities. But at the same time they were both desperately in need of new laboratory facilities. They both courted the drug industry for help, successfully in each case but with different styles.

At Oxford, head of department David Smith hosted a workshop in 1986 to demonstrate to industry what his department had to offer in terms of a basic research environment. The idea of major collaborative projects was quite new in Britain at that time and, according to Smith, many companies were unable to accept that research in the department would continue to be fundamental — and freely chosen — at all costs. But the US company Squibb (now Bristol-Myers Squibb) was in tune with this thinking. Its best-selling drug Captopril (an ACE inhibitor used in the treatment of hypertension) was a product that had emerged from a basic research approach. Captopril took 15 years from concept to market; Squibb was prepared to accept a similar timescale for discoveries in the neurosciences.

Within 12 months of the first meeting, an agreement had been signed that gave the university £20 million over seven

years to fund neuropharmacology. Half of the money was for a new, 50,000-square-foot building, the rest for research projects initiated within the department (the contract excludes research projects initiated by Squibb). Squibb has stipulated five broad areas — all within neuroscience — that it is prepared to fund. Any scientist working in these areas must apply for Squibb funding. But if the application is rejected following Squibb's peer review, money can be sought from any other source. The Squibb money constitutes about a third of all the department's outside income.

Squibb was attracted to Oxford by the concentration of neuroscience work in the area. Eight patent applications have arisen from the first five years of work, three of which, all relating to acetylcholinesterase and the diagnosis and treatment of neurological disease, have been granted. Squibb has been so pleased with the productivity that it has just agreed a further grant of £5 million to extend the collaboration for a further five years to 1999. It has, however, decided to restrict the scope of the work to Alzheimer's disease and the genetics of neurodegeneration.

Output in terms of research papers has also been impressive, with more than a hundred publications emerging from the funded work. Freedom to publish has always been a sensitive point; the